

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 13:26:20 ; Search time 25 Seconds
(without alignments)
601.931 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MRFVLVSLILTLFLTPA.....AAMDDVPMVTSFGCGSYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1422	99.3	289	1	RIP2 TRIKI
2	1377	96.2	289	1	RIPS TRIKI
3	1164.5	81.3	290	1	RIP1 BRYDI
4	859.5	60.0	286	1	RIP1 MOMCH
5	802.5	56.0	277	1	RIPA LUFY
6	756	52.8	286	1	RIP2 MOMBA
7	736	51.4	286	1	RIP1 CUCFI
8	679.5	47.5	294	1	RIP1 TRIAN
9	673	47.0	250	1	RIPB LUFY
10	537.5	37.5	282	1	RIP2 BRYDI
11	432.5	30.2	563	1	NIGB SAMNI
12	419	29.3	576	1	RIC1 RICCO
13	394.5	27.5	564	1	AGGL RICCO
14	343	24.0	562	1	ABRC ABRPR
15	337.5	23.6	527	1	ABRA ABRPR
16	333	23.3	528	1	ABRA ABRPR
17	332	23.2	254	1	MLA3 VISAL
18	322	22.5	254	1	MLA1 VISAL
19	319	22.3	316	1	RIPG GELMU
20	271.5	19.0	294	1	RIPA PHVAM
21	265	18.5	313	1	RIP1 PHVAM
22	237.5	16.6	261	1	RIPS PHVAM
23	215.5	15.0	310	1	RIP2 PHVAM
24	198.5	13.9	278	1	RIPP MIRJA
25	168	11.7	45	1	RIP2 TRIKI
26	168	11.7	299	1	RIP6 SAPOF
27	167	11.7	292	1	RIP7 SAPOF
28	162.5	11.3	253	1	RIP7 SAPOF
29	161.5	11.3	253	1	RIP5 SAPOF
30	152.5	10.6	293	1	RIP0 DIACA
31	138.5	9.7	236	1	RIP3 SAPOF
32	124.5	8.7	319	1	SLTA BP333
33	121	8.4	280	1	RIP2 HORVU

34	118.5	8.3	315	1	SLTA BPH19
35	117.5	8.2	315	1	SLTA BPH30
36	117	8.2	280	1	RIP1 HORVU
37	102.5	7.2	560	1	JI60 HORVU
38	97	6.8	454	1	DNAC BACSU
39	95.5	6.7	504	1	FLIC SALSE
40	94.5	6.6	504	1	FLIC SALBU
41	92.5	6.5	939	1	SLAP CAMPE
42	92	6.4	44	1	RIP3 MOMCH
43	92	6.4	475	1	MURC WIGBR
44	91.5	6.4	2748	1	NUM1 YEAST
45	91	6.4	468	1	HPCC ECOLI

ALIGNMENTS

RESULT 1
RIP2 TRIKI
ID RIP2 TRIKI STANDARD; PRT; 289 AA.
AC P09959;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribosome-inactivating protein alpha-trichosanthin precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase) (Alpha-TCS).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids 1; Cucurbitales; Cucurbitaceae; Trichosanthes.
ON NCBI_TaxID=3677;
RX SHAW P.C., YUNG M.H., ZHU R.H., HO W.K.K., NG T.B., YEUNG H.W.;
RT "Cloning of trichosanthin cDNA and its expression in Escherichia coli."
RL Gene 97:267-272(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz;
RX MEDLINE=91153657; PubMed=199291;
RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;
RT "Cloning of trichosanthin cDNA and its expression in Escherichia coli."
RL Gene 97:267-272(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz; TISSUE=Leaf;
RX MEDLINE=90256790; PubMed=2341400;
RA Chow T., Feldman R.A., Lovett M., Piatak M.;
RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I ribosome-inactivating protein."
RX J. Biol. Chem. 265:8670-8674 (1990).
RN [3]
RP SEQUENCE OF 24-270.
RC STRAIN=Maximowicz; TISSUE=Tubercous root;
RX MEDLINE=90256789; PubMed=2341399;
RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R., Wu P., Hwang K., Piatak M.;
RT "Primary amino acid sequence of alpha-trichosanthin and molecular models for abrin A-chain and alpha-trichosanthin."
RX J. Biol. Chem. 265:8665-8669 (1990).
RN [4]
RP SEQUENCE OF 24-270.
RC TISSUE=tubercous root;
RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X., Tian G.Y., Ni C.Z.;
RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application."
RX Pure Appl. Chem. 58:789-798 (1986).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).
RX MEDLINE=943444957; PubMed=8066085;
RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;
RT "Structure of trichosanthin at 1.88-A resolution."
RX Proteins 19:4-13(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX MEDLINE=95344383; PubMed=7619070;

P08026	bacterioph
P10149	bacterioph
P22444	hordeum vul
Q00531	hordeum vul
P37469	bacillus su
Q06983	salmonella
Q06969	salmonella
P35827	campylobac
P24817	momordica c
Q8d227	wiggleswort
Q00402	saccharomyc
P42269	escherichia

RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
RL depurinating mechanism of two ribosome-inactivating proteins.";
CC Biochem. J. 309:285-298(1995).
CC -!- FUNCTION: Trichosanthin is an abortion-inducing protein. It is
CC capable of inhibiting HIV-1 infection and replication. It
CC inactivates eukaryotic 60S ribosomal subunits.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
CC Type 1 RIP subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M34858; AAA34207.1; -;
CC EMBL; J05434; AAA34206.1; -;
CC PIR; JT0566; RLTT.
CC PDB; 1MRJ; 07-FEB-95.
CC PDB; 1MRK; 07-FEB-95.
CC PDB; 1TCS; 10-JUL-95.
CC PDB; 1J4G; 28-JAN-03.
CC PDB; 1NLI; 17-JUN-03.
CC PDB; 1QB2; 24-APR-00.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00161; RIP; 1.
CC PRINTS; PR00396; SHIGARICIN.
CC PROSITE; PS00275; SHIGA_RICIN; 1.
CC Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 270
FT
FT PROPEP 271 289
FT ACT SITE 183 183
FT CONFLICT 57 60
FT CONFLICT 82 84
FT CONFLICT 87 87
FT CONFLICT 92 92
FT CONFLICT 143 144
FT CONFLICT 196 196
FT CONFLICT 215 216
FT CONFLICT 231 231
FT CONFLICT 234 234
FT CONFLICT 246 266
FT CONFLICT 247 247
FT STRAND 25 28
FT TURN 30 31
FT TURN 34 46
FT TURN 47 47
FT STRAND 50 54
FT TURN 55 56
FT STRAND 57 60
FT HELIX 66 69
FT STRAND 70 76
FT TURN 78 79
FT STRAND 82 88
FT TURN 89 92
FT STRAND 93 99
FT TURN 100 101
FT STRAND 102 105
FT TURN 109 114
FT TURN 115 117
FT TURN 120 121
FT STRAND 124 127
FT HELIX 134 141
FT TURN 142 142
FT HELIX 145 147

FT STRAND 150 150
FT HELIX 152 163
FT TURN 164 185
FT HELIX 167 180
FT TURN 181 181
FT HELIX 182 186
FT STRAND 187 187
FT HELIX 188 195
FT TURN 196 196
FT STRAND 202 202
FT HELIX 206 226
FT TURN 227 230
FT STRAND 231 239
FT TURN 241 242
FT STRAND 245 250
FT TURN 251 252
FT HELIX 254 258
FT TURN 259 259
FT STRAND 260 260
FT TURN 263 263
FT TURN 266 268
SQ SEQUENCE 289 AA; 31676 MW; 5CE09BB630575BB9 CRC64;
Query Match 99.3%; Score 1422; DB 1; Length 289;
Best Local Similarity 99.3%; Pred. No. 7.9e-105;
Matches 287; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MIRFLVLSLLILTLFTTPPAVEGDVSVFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRFLVLSLLILTLFTTPPAVEGDVSVFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120
DB 61 RSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDPAITTLFYNNANSAASALMVLIO 180
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDPAITTLFYNNANSAASALMVLIO 180
QY 181 TSEAAARYKFEIQIGKRVKDTLPALAIISLNSWSALSQKQIASTNNGQFSPVVLIN 240
DB 181 TSEAAARYKFEIQIGKRVKDTLPALAIISLNSWSALSQKQIASTNNGQFSPVVLIN 240
QY 241 AQNRVTIINVDAGVVTNSIALLLNNMAAMDDVPMTQSFQCGSYAL 289
DB 241 AQNRVTIINVDAGVVTNSIALLLNNMAAMDDVPMTQSFQCGSYAI 289
RESULT 2
RIPS TRIKI STANDARD; PRT; 289 AA.
AC P24478;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribosome-inactivating protein karasurin precursor (EC 3.2.2.22) (rRNA
DE N-glycosidase).
OS Trichosanthin kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid I; Cucurbitales; Cucurbitaceae; Trichosanthes.
ON NCBI_TaxID=3677;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root tuber;
RX MEDLINE=97356562; PubMed=9212998;
RA Mizukami H., Iida K., Kondo T., Ogiwara Y.;
RT "Cloning and bacterial expression of a gene encoding ribosome-
RT inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes
RT kirilowii var. japonica";
RL Biol. Pharm. Bull. 20:711-713(1997).
RN [2]
RP SEQUENCE OF 24-270.

```

RX MEDLINE=92005921; PubMed=1914000;
RA Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;
RT "The complete amino acid sequence of an abortifacient protein,
RL karasurin.";
RL Chem. Pharm. Bull. 39:1244-1249(1991).
CC -!- FUNCTION: Abortion-inducing protein. It inactivates eukaryotic
CC 60S ribosomal subunits.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
CC Type 1 RIP subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB000666; BAA21786.1; -
DR PIR; JC5606; JC5606.
DR J00393; J00393.
DR HSP; P09389; IMRJ.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR003396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 270 KARASURIN-C.
FT CHAIN 24 270 KARASURIN-A.
FT PROPEP 271 289 REMOVED IN MATURE FORM.
FT ACT SITE 183 183 BY SIMILARITY.
SQ SEQUENCE 289 AA; 31704 MW; 883D3E3242887B26 CRC64;

Query Match 96.2%; Score 1377; DB 1; Length 289;
Best Local Similarity 95.8%; Pred No. 2.7e-101;
Matches 277; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MIRFLVSLILTLFTTPAVGDSVFRSLGATSSYGVFISNLRKALPNERKLYDIPLL 60
Db 1 MIRFLVFSLLILTLFTTPAVGDSVFRSLGATSSYGVFISNLRKALPYERKLYDIPLL 60
QY 61 RSLFPGSQRYALIHLYTNYADETISVAIDVTNYVMYGRAGDTSYFPNGASATEAAKYVEK 120
Db 61 RSTLFGSQRYALIHLYTNYADETISVAIDVTNYVMYGRAGDTSYFPNEASATEAAKYVEK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAALMVLIOQ 180
Db 121 DAKRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAALMVLIOQ 180
QY 181 TSEAAKYFIEQIGKRVKDTFLPSLAISLNSALSALSKQIQIASTNNQFESPVLIN 240
Db 181 TSEAAKYFIEQIGKRVKDTFLPSLAISLNSALSALSKQIQIASTNNQFETPVLIN 240
QY 241 AQQRVTITNDAGVTSNIALLNENNAAMDDVPMTQSGCGSYAL 289
Db 241 AQQRVTITNDAGVTSNIALLNENNAAMDDVPMTQSGCGSYAI 289

RESULT 3
RIP1_BRYDI
ID RIP1_BRYDI STANDARD; PRT; 290 AA.
AC P33185; Q93819;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribosome-inactivating protein bryodin I precursor (EC 3.2.2.22) (rRNA
DE N-glycosidase) (BD1).
OS Bryonia dioica (Red Bryony).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosida I; Cucurbitales; Cucurbitaceae; Bryonia.
OX NCBI_TaxID=3652;
RN [1]
RP SEQUENCE FROM N.A., MUTAGENESIS OF GLU-12, AND X-RAY CRYSTALLOGRAPHY
RP (2.1 ANGSTROMS).
RX TISSUE=Leaf;
RC MEDLINE=97228081; PubMed=9115985;
RA Gawlak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,
RA Siegall C.B.;
RT "Molecular, biological, and preliminary structural analysis of
RT recombinant bryodin 1, a ribosome-inactivating protein from the plant
RL Bryonia dioica.";
RL Biochemistry 36:3095-3103(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Siegall C.B.;
RT "Cloning and expression of a gene encoding bryodin 1 from Bryonia
RT dioica.";
RT Patent number US5541110, 30-JUL-1996.
RN [3]
RP SEQUENCE OF 24-66.
RC TISSUE=Seed;
RX MEDLINE=89326691; PubMed=2753596;
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Sorcia M.,
RA Lappi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins.";
RL Int. J. Pept. Protein Res. 33:263-267(1989).
RN [4]
RP SEQUENCE OF 24-43.
RC TISSUE=Root;
RX MEDLINE=95151812; PubMed=7849072;
RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
RA Marquardt H.;
RT "Characterization of ribosome-inactivating proteins isolated from
RT Bryonia dioica and their utility as carcinoma-reactive
RT immunokonjugates.";
RL Bioconj. Chem. 5:423-429(1994).
CC -!- FUNCTION: Ribosome-inactivating protein of type 1, inhibits
CC protein synthesis in animal cells.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- PTM: Appears to undergo proteolytic cleavage in the C-terminal to
CC produce a shorter protein.
CC -!- BIOTECHNOLOGY: Especially useful as immunotoxin for
CC pharmacological applications as it has low toxicity in rats and
CC mice but is potent once inside target cells.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
CC Type 1 RIP subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J24020; -; NOT ANNOTATED_CDS.
DR PIR; S16491; S16491.
DR PDB; 1BRY; 04-MAR-98.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR003396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
KW 3D-structure; Multigene family; Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 270 RIBOSOME-INACTIVATING PROTEIN BRYODIN 1.
FT PROPEP 271 290 MISSING IN MATURE PROTEIN.
FT ACT SITE 183 183 BY SIMILARITY.
FT ACT SITE 212 212
FT CARBOHYD 214 214 N-LINKED (GLCNAC...) (POTENTIAL).

```

FT CARBOHYD 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MUTAGEN 212 E->K: 10-FOLD REDUCTION IN ACTIVITY.
 FT CONFLICT 61 RSSIS -> LRHXI (IN REF. 3).
 FT STRAND 25
 FT TURN 30 31
 FT HELIX 34 46
 FT TURN 47 47
 FT STRAND 50 54
 FT TURN 55 56
 FT STRAND 57 60
 FT HELIX 66 69
 FT STRAND 70 76
 FT TURN 78 79
 FT STRAND 82 88
 FT TURN 89 92
 FT STRAND 93 99
 FT TURN 100 101
 FT STRAND 102 105
 FT HELIX 109 114
 FT TURN 115 117
 FT TURN 120 121
 FT STRAND 124 127
 FT HELIX 134 141
 FT TURN 142 142
 FT STRAND 145 147
 FT STRAND 150 150
 FT HELIX 152 163
 FT TURN 164 165
 FT HELIX 167 186
 FT STRAND 187 187
 FT HELIX 188 196
 FT STRAND 202 202
 FT HELIX 206 213
 FT TURN 214 214
 FT HELIX 215 225
 FT TURN 226 230
 FT STRAND 231 239
 FT TURN 241 242
 FT STRAND 245 250
 FT TURN 251 252
 FT HELIX 254 257
 FT TURN 258 259
 FT STRAND 260 260
 FT STRAND 263 263
 FT HELIX 266 268
 SQ SEQUENCE 290 AA; 31788 MW; E966CD9C031A42DB CRC64;

Query Match 81.3%; Score 1164.5; DB 1; Length 290;
 Best Local Similarity 81.0%; Pred. No. 1.5e-84;
 Matches 235; Conservative 26; Mismatches 28; Indels 1; Gaps 1;

QY 1 MIRELVLSLILTLFTTTPAVSGDVSFRISGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 Db 1 MKLVLLWLLIITFLKSTVSGDVSFRISGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPQSORYALIHLYNVADEITISVAIDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVEK 120
 Db 61 RSSISGSGRYTLHLITNVADEITISVAVDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVEK 120

QY 121 DAMKVTPLPSGNYERLQTAACKIRENIPGLPALDSAITTLFYNNANSASALMVLIQS 180
 Db 121 DAKKVTPLPSGNYERLQTAACKIRENIPGLPALDSAITTLFYNNANSASALMVLIQS 180

QY 181 TSEARYKFIEQIGKRVKDTPLSLAIIISLNSWSLSKQIQIASTNNNGOPESPVLIN 240
 Db 181 TAESARYKFIEQIGKRVKDTPLSLAIIISLNSWSLSKQIQIASTNNNGOPESPVLID 240

QY 241 AQNRVTITNDAGVVTNSIALLLNRNNNAAMDDDPMTQ-SFGCGSYAL 289
 Db 241 GNNQVRSITNASARVVTNSIALLLNRNNNAIAGEDISMTLIGFERGLYGI 290

RIPI1 MOMCH
 ID RIPI1 MOMCH STANDARD; PRT; 286 AA.
 AC P16094; P24697;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ribosome-inactivating protein momordin I precursor (EC 3.2.2.22) (rRNA
 DE N-glycosidase) (Alpha-momorcharin) (Alpha-WMC)
 OS Momordica charantia (Bitter melon) (Balsam pear).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eusoids I; Cucurbitales; Cucurbitaceae; Momordica.
 OC NCBI_TaxID=3673;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Seed;
 RX MEDLINE=91159486; PubMed=2001404;
 RA Ho W.K.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
 RT "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating
 RT protein.";
 RL Biochim. Biophys. Acta 1088:311-314(1991).
 RN [2]
 RP SEQUENCE OF 24-38.
 RC TISSUE=Seed;
 RX MEDLINE=89326691; PubMed=2753596;
 RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
 RA Lappi D.;
 RT "N-terminal sequence of some ribosome-inactivating proteins.";
 RL Int. J. Pept. Protein Res. 33:263-267(1989).
 RN [3]
 RP SEQUENCE OF 24-70.
 RC TISSUE=Seed;
 RX MEDLINE=89005108; PubMed=3262509;
 RA Casellas P., Dussosoy D., Falasca A.I., Barbieri L., Stirpe F.;
 RA Guillemot J.C., Ferrara P., Bolognesi A., Cennini P., Stirpe F.;
 RT "Trichokirin, a ribosome-inactivating protein from the seeds of
 RT Trichosanthes kirilowii Maximowicz. Purification, partial
 RT characterization and use for preparation of immunotoxins.";
 RL Eur. J. Biochem. 176:581-588(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=94356447; PubMed=8075985;
 RA Ren J., Wang Y., Dong Y., Stuart D.I.;
 RT "The N-glycosidase mechanism of ribosome-inactivating proteins
 RT implied by crystal structures of alpha-momorcharin.";
 RL Structure 2:7-16(1994).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
 RX MEDLINE=94192822; PubMed=8143869;
 RA Huaain J., Tickle I.J., Wood S.P.;
 RT "Crystal structure of momordin, a type I ribosome inactivating
 RT protein from the seeds of Momordica charantia.";
 RL FEBS Lett. 342:154-158(1994).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=95344383; PubMed=7619070;
 RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
 RT "Studies on crystal structures, active-centre geometry and
 RT depurinating mechanism of two ribosome-inactivating proteins.";
 RL Biochem. J. 309:285-298(1995).
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
 CC Type 1 Rlp subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

Db 61 LPSVSGAGRYLLMHLFNYDGKTTTAVDVNTVNGVLADTTSYFFNEPRAELASQVYFR 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLGLPALDSAITTLFFYNANSASALMWLIQS 180

Db 121 DARRKTTLPYSGNYERLQTAAGKPREKIPGLGLPALDSAITLLHYDSTAAGALLVLIQT 180

QY 181 TSPAARYKFTIEQOIGKVDKFTPLPSLAISLNSWSALSQIQIAGTNNQFSPVVLIN 240

Db 181 TAEAAFKFTIEQIOERAYRDEVPSLATISLNSWSGLSQIQIOLAQNGNGIFRTPIVLND 240

QY 241 AQNQRVTITNDVAGVTSNIALLNENNAAMD-DVPMQTQSF 282

Db 241 NKNRNVQITNVTSKVTSNIQLLNTNTNIAEGONGDVSTHGF 283

RESULT 5
RIPA_LUFFCY
ID RIPA_LUFFCY STANDARD; PRT; 277 AA.
AC Q00455; 1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribosome-inactivating protein luffin-alpha precursor (EC 3.2.2.22)
DE (rRNA N-glycosidase)
OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.
NCBI_TaxID=3670;
[1]
SEQUENCE FROM N.A.
RN RNP
RP TISSUE=Seed.
RC MEDLINE=92288316; PubMed=1600156;
RA Kataoka J., Habuka N., Miyano M., Masuta C., Koiwai A.;
RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
RT inactivating protein from Luffa cylindrica.";
RL Plant Mol. Biol. 18:1199-1202 (1992).
CC -|- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -|- SIMILARITY: Belongs to the ribosome-inactivating protein family.
CC Type 1 RIP subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X62371; CAA44229.1; --
DR PIR; S22494; S22494.
DR HSSP; P16094; 1AHC.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal.
FT SIGNAL 1 19
FT CHAIN 20 277 RIBOSOME-INACTIVATING PROTEIN LUFFIN-
FT ALPHA.
FT ACT SITE 179 179 BY SIMILARITY.
FT SEQUENCE 277 AA; 30212 MW; EAL7FC27998C25AC CRC64;
Query Match 56.0%; Score 802.5; DB 1; Length 277;
Best Local Similarity 60.3%; Pred. No. 4.3e-56;
Matches 167; Conservative 45; Mismatches 60; Indels 5; Gaps 3;
QY 1 MTRFVLVLSLILTLFTTAVGDSVFRISGATSSSYGVFISNRKALPNERKLYDIPLL 60
Db 1 MKRFVTV---LILAIFVAASTVEADVRFSSGSSSTSYKFIQDLRKALPNSGVYNTLL 57


```

Db      1  MNRSLVCLVLSIFGHGPTAAGDVTKFSLGSHKHSYSKFITSMNALPNACDIYNI 60
      58  PLLSSLPQSORYALIHITNVADETISVAIDVTNVMYIMGYRAGDTSYFNGASATEAAKY 117
      61  PLLVPSISGSRRIYLMQLSNYEGNTITMAVDVTNVMYIMGYLVNGTSYFNETDAQLASKF 120
      118  VFQ-OTKSTITLPSYNGYERLQTAAGKIRENIPGLPALDSATITLFFYNANSAAALMVL 177
      121  VFQ-OTKSTITLPSYNGYERLQTAAGKIRENIPGLPALDSATITLFFYNANSAAALMVL 179
      178  IQSTSEAAARYKFEIOQIGKRVDTKFLPSLAISLNSWSALSQKIQIASTNNQGFESPVV 237
      180  IOTTAEAARYKYIEQIIDRISVSKVPDLAALSLENWSLLSKQIQIAKSNNGQFQTPVK 239
      238  LINAQQRVTITNDVAGVTSNIALLNRRNNMAAMDDVPMT 279
      240  IINDKGILTEVNTVSSLVVTKNIMLLNKLNTASFEDHVIST 281

```

RESULT 8

```

ID RIPI TRIAN STANDARD; PRT; 294 AA.
AC P56626; Q9ZQY7;
DT 15-DEC-1998 (Rel. 37, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Type 1 ribosome-inactivating protein trichoanguina precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase) (RIP) (Trichoanguin).
GN TCA.
OS Trichosanthes anguina (Snake gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=50544;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC STRAIN=cv. Anguina; TISSUE=Seed;
RX MEDLINE=99132006; PubMed=9931318;
RA Chow L.-P., Chou M.-H., Ho C.-Y., Chuang C.-C., Pan F.-M., Wu S.-H.,
RA Lin J.-Y.;
RT "Purification, characterization and molecular cloning of trichoanguin,
RT a novel type I ribosome-inactivating protein from the seeds of
RT Trichosanthes anguina."
RL Biochem. J. 338:211-219(1999).
RN [2]
RP SEQUENCE OF 20-264.
RC TISSUE=Seed;
RA Chow L.-P., Kamo M., Lin J.-Y., Wang S.-H., Ueno Y., Tsugita A.;
RT "Amino acid sequence of trichoanguina, a ribosomal-inactivating
RT protein from Trichosanthes anguina seeds."
RL J. Biomed. Sci. 3:178-186(1996).
CC -!- FUNCTION: Inhibits protein synthesis by depurinating 28S rRNA in
CC ribosomes.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
CC Type 1 RIP subfamily.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; AF055086; RAD02686.1; -
CC HSSP; P33185; LBRY.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00161; RIP; 1.
CC PRINTS; PR00396; SHIGARICIN.
CC PROSITE; PS00275; SHIGA_RICIN; FALSE_NEG.

```

```

KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
KW Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 264
      TYPE I RIBOSOME-INACTIVATING PROTEIN
      TRICHOANGUINA.
FT PROPEP 265 294
FT ACT_SITE 177 177
FT ACT_SITE 180 180
FT CARBOHYD 70 70
FT CARBOHYD 220 220
FT CONFLICT 51 51
FT CONFLICT 65 65
FT CONFLICT 84 84
FT CONFLICT 152 152
FT CONFLICT 174 174
FT CONFLICT 245 245
FT SEQUENCE 294 AA; 32234 MW; DA4F8B7CE3290994 CRC64;
      Query Match 47.5%; Score 679.5; DB 1; Length 294;
      Best Local Similarity 52.2%; Pred. No. 2.3e-46;
      Matches 144; Conservative 38; Mismatches 91; Indels 3; Gaps 3;
QY 5 LVLSLLILTLFTTAVEGDVSRFSGATSSSYGVFISNLKALPNERKLYDIPILRSGL 64
Db 1 MALSPFFLAISLGSPTAIGDVSFDSLSTATKXSYSSFTQLRDALPTQGTVCIGILLPSTA 60
QY 65 PGSQRYALIHITNVADETISVAIDVTNVMYIMGYRAGDTSYFNGASATEAAKYVFKDAMR 124
Db 61 SGSQWFRFNLNTNVDVAGVTSNIALLNRRNNMAAMDDVPMT 279
QY 125 KVTLPYSNGYERLQTAAGKIRENIPGLPALDSATITLFFYNANSAAALMVLIOSTSEA 184
Db 119 TVKLPYSNGYDKLQSVGKQRMIELGIPALSSAITNNVYDYQSTAAALLVLIQCTAEA 178
QY 185 ARYKFEIOQIGKRVDTKFLPSLAISLNSWSALSQKIQIAS-TNNGQFESPVLINAQ 243
Db 179 ARYKFEIOQVSSHSSHNSFNPQAVISLENKWLKSKQIQIANRTGHGQFENPVELNPDG 238
QY 244 QRVTITNDVAGVTSNIALLNRRNNMAAMDDVPMT 279
Db 239 TRFSVTNTSAGVVGKGNIKLLLYKASGVSEYDIPTT 274
      RESULT 9
      RIPB_LUFUCY STANDARD; PRT; 250 AA.
      ID RIPB_LUFUCY STANDARD; PRT; 250 AA.
      AC P22851;
      DT 01-AUG-1991 (Rel. 19, Created)
      DT 01-AUG-1991 (Rel. 19, Last sequence update)
      DT 15-MAR-2004 (Rel. 43, Last annotation update)
      DE Ribosome-inactivating protein luffin-B (EC 3.2.2.22) (rRNA N-
      DE glycosidase).
      OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
      OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
      OC eucosids I; Cucurbitales; Cucurbitaceae; Luffa.
      OX NCBI_TaxID=3670;
      RN [1]
      RP SEQUENCE.
      RC TISSUE=Seed;
      RX MEDLINE=91248488; PubMed=1368666;
      RA Islam M.R., Hirayama H., Funatsu G.;
      RT "Complete amino acid sequence of luffin-b, a ribosome-inactivating
      RT protein from sponge gourd (luffa cylindrica) seeds."
      RL Agric. Biol. Chem. 55:229-238(1991).
      CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
      CC specific adenosine on the 28S rRNA.
      CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
      CC Type 1 RIP subfamily.
      CC PIR; JN0108; JN0108.
      CC HSSP; P16094; 1AHC.
      CC InterPro; IPR001574; RIP.
      CC Pfam; PF00161; RIP; 1.

```



```
Query Match      29.3%; Score 419; DB 1; Length 576;
Best Local Similarity 36.6%; Pred. No. 1.7e-25;
Matches 93; Conservative 60; Mismatches 85; Indels 16; Gaps 7;

QY 25 VSFRUSGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSSLPQSORYALHILHTNYADE 81
Dd 44 INFTTAGATVQSYTTFIRAVRGLTTGADVREIPLVLPNRVGLPIQRFILVELSNHAE 103

QY 82 TISVALDVTNVIMGYRAGDTSYFF--NGASATEAAKYVFKDAMRKVTLPLYSNVERLOT 139
Dd 104 SVTLALDVTNAVYVGRAGNSAYFFHPDQDAEAITHLFTDVQNRYPFAFGNYDRLEQ 163

QY 140 AAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMVLIOSTSEAAKYKIEQOI 194
Dd 164 LAGNURENIELNGPLEBAISALYYSTGTOPLTLARSFICIQMISEAARFQIEGEM 223

QY 195 GKR--DKTFLPSLAISLNSWALSLSKQIQIASTNNQFSPVVLINQONRVITNVD 252
Dd 224 RTRIRYRRSADPSPVITLNSWGRSLTAIQ--ESNQAFASPIQLQRNGSKFSV--YD 279

QY 253 AGVTSNTALLNR 266
Dd 280 VSILIPITALLMVR 293

RESULT 13
AGGL_RICCO
ID AGGL_RICCO STANDARD; PRT; 564 AA.
AC P06750;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (EC 3.2.2.22)
DE (rRNA N-glycosidase); Agglutinin B chain].
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaceae;
OC Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86059449; PubMed=2999130;
RA Roberts L.M., Lamb F.L., Pappin D.J.C., Lord J.M.;
RT "The primary sequence of Ricinus communis agglutinin. Comparison with
RT ricin.";
RL J. Biol. Chem. 260:15682-15686(1985).
RN [2]
RP SEQUENCE OF 303-564.
RC TISSUE=Seed;
RA Araki T., Yoshioka Y., Funatsu G.;
RT "The complete amino acid sequence of the B-chain of the Ricinus
RL communis agglutinin isolated from large-grain castor bean seeds.";
RL Biochim. Biophys. Acta 872:277-285(1986).
RN [3]
RP SEQUENCE OF 303-337.
RX MEDLINE=80178723; PubMed=6768555;
RA Lin T.-S., Li S.-L.;
RT "Purification and physicochemical properties of ricins and
RL agglutinins from Ricinus communis.";
RL Eur. J. Biochem. 105:453-459(1980).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: In the N-terminal section; belongs to the ribosome-
CC inactivating protein family. Type 2 RIP subfamily.
CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
Dd EMBL; M12089; AAA33869.1; -
Dd EMBL; S40368; AAB22584.1; -
Dd PIR; A24261; RLCSAG.
Dd HSSP; P02879; 1BR6.
Dd GlycoSuiteDB; P06750; -
Dd InterPro; IPR000772; Ricin_B_lectin.
Dd InterPro; IPR008997; RicinB_Like.
Dd InterPro; IPR001574; RIP.
Dd Pfam; PF00652; Ricin_B_lectin; 6.
Dd Pfam; PF00161; RIP; 1.
Dd PRINTS; PRO00396; SHIGARICIN.
Dd SMART; SM00458; RICIN; 2.
Dd PROSITE; PS50231; RICIN_B_LECTIN; 2.
Dd PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; Signal.
FT SIGNAL 1 24
FT CHAIN 25 290 AGGLUTININ A CHAIN.
FT PROPEP 291 302 LINKER PEPTIDE.
FT CHAIN 303 564 AGGLUTININ B CHAIN.
FT DOMAIN 309 436 RICIN B-TYPE LECTIN 1.
FT DOMAIN 439 563 RICIN B-TYPE LECTIN 2.
FT REPEAT 319 361 1-ALPHA.
FT REPEAT 362 402 1-BETA.
FT REPEAT 405 437 1-GAMMA.
FT REPEAT 450 485 2-ALPHA.
FT REPEAT 489 528 2-BETA.
FT REPEAT 531 558 2-GAMMA.
FT ACT_SITE 200 200 BY SIMILARITY.
FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).
FT DISULFID 322 341 BY SIMILARITY.
FT DISULFID 385 382 BY SIMILARITY.
FT DISULFID 453 466 BY SIMILARITY.
FT DISULFID 492 509 BY SIMILARITY.
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 337 397 N-LINKED (GLCNAC. .).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .).
FT CONFLICT 331 331 F -> T (IN REF. 2).
FT CONFLICT 362 362 N -> D (IN REF. 2).
FT CONFLICT 374 374 R -> G (IN REF. 2).
FT CONFLICT 404 404 R -> T (IN REF. 2).
FT CONFLICT 552 552 F -> V (IN REF. 2).
SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CRC64;

Query Match      27.5%; Score 394.5; DB 1; Length 564;
Best Local Similarity 36.6%; Pred. No. 1.4e-23;
Matches 93; Conservative 55; Mismatches 89; Indels 17; Gaps 8;

QY 25 VSFRUSGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSSLPQSORYALHILHTNYADE 81
Dd 33 INFTTAGATVQSYTTFIRAVRGLTTGADVREIPLVLPNRVGLPIQRFILVELSNHAE 92

QY 82 TISVALDVTNVIMGYRAGDTSYFF--NGASATEAAKYVFKDAMRKVTLPLYSNVERLOT 139
Dd 93 SVTLALDVTNAVYVGRAGNSAYFFHPDQDAEAITHLFTDVQNSFTFAFGNYDRLEQ 152

QY 140 AAGKIRENIPGLPALDSAITTLFYNNAN-----NSAASALMVLIOSTSEAAKYKIEQOI 194
Dd 153 LGG-LRENIELGTGPLEDAISALYYSTGTOPLTLARSFVCIQMWISEAARFQIEGEM 211

QY 195 GKR--DKTFLPSLAISLNSWALSLSKQIQIASTNNQFSPVVLINQONRVITNVD 252
Dd 212 RTRIRYRRSADPSPVITLNSWGRSLTAIQ--ESNQAFASPIQLQRNGSKFSV--YD 267

QY 253 AGVTSNTALLNR 266
Dd 269 VSILIPITALLMVR 281
```

```

RESULT 14
ABRC ABRPR STANDARD; PRT; 562 AA.
ID ABRP ABRPR STANDARD; PRT; 527 AA.
AC Q06077; P81374;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ABRin-b precursor [Contains: ABRin-b A chain (EC 3.2.2.22) (rRNA N-glycosidase); ABRin-b B chain].
OS ABRus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=931266957; PubMed=2050149;
RA Wood K.A., Lord J.M., Wawrzyniak E.J., Piatak M.;
RT "Preproabrin: genomic cloning, characterisation and the expression of
RT the A-chain in Escherichia coli.";
RL Eur. J. Biochem. 198;723-732(1991).
CC -!- FUNCTION: The A chain is responsible for inhibiting protein
CC synthesis through the catalytic inactivation of 60S ribosomal
CC subunits by removing adenine from position 4,324 of 28S rRNA.
CC ABRin-a is more toxic than ricin.
CC -!- FUNCTION: The B chain is a galactose-specific lectin that
CC facilitates the binding of abrin to the cell membrane that
CC precedes endocytosis.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SUBUNIT: Disulfide-linked dimer of A and B chains.
CC -!- DOMAIN: The B chain is composed of two domains, each domain
CC consists of 3 homologous subdomains (alpha, beta, gamma).
CC -!- SIMILARITY: In the N-terminal section; belongs to the ribosome-
CC inactivating protein family. Type 2 RIP subfamily.
CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X55667; CA319202.1;
CC PIR; S16022; S16022.
CC HSP; F11140; IABR.
CC InterPro; IPR000772; Ricin_B_lectin.
CC InterPro; IPR008997; RicinB_like.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00652; Ricin_B_lectin; 6.
CC Pfam; PF00161; RIP; I_lectin.
CC PRINTS; PR00396; SHIGARICIN.
CC SMART; SM00458; RICIN; 2.
CC PROSITE; PS00231; RICIN_B_LECTIN; 2.
CC PROSITE; PS00275; SHIGA_RICIN; 1.
CC Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
CC Glycoprotein; Lectin; Signal; Pyrrolidone carboxylic acid.
CC SIGNAL 1 34
CC CHAIN 35 285
CC PEPTIDE 286 295
CC CHAIN 296 562
CC DOMAIN 307 434
CC DOMAIN 437 561
CC REPEAT 317 359
CC REPEAT 360 400
CC REPEAT 403 435
CC REPEAT 448 483
CC REPEAT 487 526
CC REPEAT 529 562
CC ACT_SITE 198
CC BY SIMILARITY.

FT DISULFID 281 303 INTERCHAIN (BY SIMILARITY).
FT DISULFID 320 339 BY SIMILARITY.
FT DISULFID 363 380 BY SIMILARITY.
FT DISULFID 451 464 BY SIMILARITY.
FT DISULFID 490 507 BY SIMILARITY.
FT MOD_RES 35 PYRROLIDONE CARBOXYLIC ACID
FT CARBOHYD 234 234 (BY SIMILARITY).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 562 AA; 62817 MW; 1FD0ABC7D7BA6278 CRC64;
Query Match 24.0%; Score 343; DB 1; Length 562;
Best Local Similarity 36.4%; Pred. No. 1.6e-19;
Matches 102; Conservative 43; Mismatches 107; Indels 28; Gaps 10;
QY 7 LSLILITLFLTT-----PAVEGD-----VSFRLSGATSSSYGVFTSNLRKALPN 50
DB 5 LKLLILCLAWTCFSALRCAARTYPPVATNQDVIKFTTEGATSSQSYKQFTEALQRUTG 64
QY 51 ERKYLDIPLL--RSSLPQSQRVALIHLTNVADETISVAIDVTNVYIMGVYAGDTSYFFNG 108
DB 65 -GLIHDIPLVDPPTVEERNRYITVELSNRESIEVGIDVTNAYVAYRAGSQSYFLRD 123
QY 109 ASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTL--FYNN 166
DB 124 APAS-ASTYLPFGTQR-YSLRFDGSGYGLERWAHQTREISLGLQALTHAISFLRSGASN 181
QY 167 ANSASALMVLIOSTSEAAKYKFEQOIGK--RVDKTPSLSLAIISLNSWSLSKQIQI 224
DB 182 DEEKARTLIVTIQWASEAARYISNRVGSIRGTAFQPPAPMLSLNNNDNLSSGGVQ- 240
QY 225 ASTNNGQFSPVVLINAGNQRTVTNVDAGVVTNSIALLL 264
DB 241 -QSVQDTFNNVILLSSINRQPVVDLSLHPTV-AVLALML 278

RESULT 15
ABRB ABRPR STANDARD; PRT; 527 AA.
ID ABRB ABRPR STANDARD; PRT; 527 AA.
AC Q06077; P81374;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ABRin-b precursor [Contains: ABRin-b A chain (EC 3.2.2.22) (rRNA N-glycosidase); ABRin-b B chain].
OS ABRus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93132798; PubMed=8421313;
RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
RT "Primary structure of three distinct isoabrin determined by cDNA
RT sequencing. Conservation and significance.";
RL J. Mol. Biol. 229:263-267(1993).
RN [2]
RP SEQUENCE OF 260-527.
RC TISSUE=Seed;
RX MEDLINE=93169023; PubMed=7763422;
RA Kimura M., Sumizawa T., Funatsu G.;
RT "The complete amino acid sequences of the B-chains of abrin-a and
RT abrin-b, toxic proteins from the seeds of Abrus precatorius.";
RL Biosci. Biotechnol. Biochem. 57:166-169(1993).
CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RRNA.
CC ABRIN-A IS MORE TOXIC THAN RICIN.
CC -!- FUNCTION: The B chain is a galactose-specific lectin that
CC facilitates the binding of abrin to the cell membrane that
CC precedes endocytosis.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 13:56:41 ; Search time 33 Seconds
(without alignments)
452.118 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MIRFLVLSLILTLFLTPA.....AAMDDVPMTQSGCGSYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1413	98.7	289	1 US-07-923-692C-4	Sequence 4, Appli
2	1413	98.7	289	1 US-08-184-237-4	Sequence 4, Appli
3	1413	98.7	289	1 US-08-482-920-4	Sequence 4, Appli
4	1413	98.7	289	3 US-08-484-341-4	Sequence 4, Appli
5	1413	98.7	289	3 US-08-483-502-4	Sequence 4, Appli
6	1413	98.7	289	4 US-09-726-651A-4	Sequence 4, Appli
7	1321	92.2	267	1 US-08-378-761A-74	Sequence 74, Appl
8	1321	92.2	267	1 US-08-485-286-74	Sequence 74, Appl
9	1202	83.9	247	1 US-08-488-113B-6	Sequence 6, Appli
10	1202	83.9	247	1 US-08-477-484B-6	Sequence 6, Appli
11	1202	83.9	247	2 US-08-646-360-6	Sequence 6, Appli
12	1202	83.9	247	3 US-08-839-765-6	Sequence 6, Appli
13	1202	83.9	247	3 US-09-136-389-6	Sequence 6, Appli
14	1202	83.9	247	4 US-09-610-838-6	Sequence 6, Appli
15	1202	83.9	247	4 US-09-711-485-6	Sequence 6, Appli
16	1164.5	81.3	290	1 US-08-245-754A-2	Sequence 2, Appli
17	1164.5	81.3	290	2 US-08-597-731-2	Sequence 2, Appli
18	1130	78.9	255	1 US-07-901-707-6	Sequence 6, Appli
19	1130	78.9	255	1 US-07-988-430-6	Sequence 6, Appli
20	1130	78.9	255	1 US-08-425-336-6	Sequence 6, Appli
21	1130	78.9	255	5 PCT-US92-09487-6	Sequence 6, Appli
22	1074	75.0	496	3 US-08-902-486-15	Sequence 15, Appl
23	1064	74.3	248	3 US-08-902-486-7	Sequence 7, Appli
24	806.5	56.3	263	1 US-07-901-707-7	Sequence 7, Appli
25	806.5	56.3	263	1 US-07-988-430-7	Sequence 7, Appli
26	806.5	56.3	263	1 US-08-425-336-7	Sequence 7, Appli
27	806.5	56.3	263	1 US-08-488-113B-7	Sequence 7, Appli

28	806.5	56.3	263	1 US-08-477-484B-7	Sequence 7, Appli
29	806.5	56.3	263	2 US-08-646-360-7	Sequence 7, Appli
30	806.5	56.3	263	3 US-08-839-765-7	Sequence 7, Appli
31	806.5	56.3	263	3 US-09-136-389-7	Sequence 7, Appli
32	806.5	56.3	263	4 US-09-610-838-7	Sequence 7, Appli
33	806.5	56.3	263	4 US-09-711-485-7	Sequence 7, Appli
34	806.5	56.3	263	5 PCT-US92-09487-7	Sequence 7, Appli
35	760	53.1	250	1 US-08-378-761A-78	Sequence 78, Appl
36	760	53.1	250	1 US-08-485-286-78	Sequence 78, Appl
37	751	52.4	286	1 US-08-324-301-13	Sequence 13, Appl
38	717.5	50.1	248	1 US-07-901-707-5	Sequence 5, Appli
39	717.5	50.1	248	1 US-07-988-430-5	Sequence 5, Appli
40	717.5	50.1	248	1 US-08-425-336-5	Sequence 5, Appli
41	717.5	50.1	248	1 US-08-378-761A-75	Sequence 75, Appl
42	717.5	50.1	248	1 US-08-485-286-75	Sequence 75, Appl
43	717.5	50.1	248	1 US-08-488-113B-5	Sequence 5, Appli
44	717.5	50.1	248	1 US-08-477-484B-5	Sequence 5, Appli
45	717.5	50.1	248	2 US-08-646-360-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-07-923-692C-4
; Sequence 4, Application US/07923692C
; Patent No. 5316931
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantham, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann Myers
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CAL
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,692C
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 600,244
; FILING DATE: 22-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 641,617
; FILING DATE: 16-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 310,881
; FILING DATE: 17-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 160,766
; FILING DATE: 26-FEB-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 160,771
; FILING DATE: 26-FEB-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 347,637
; FILING DATE: 05-MAY-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 363,138
; FILING DATE: 08-JUN-1989

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 219,279
; FILING DATE: 15-JUL-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: BIOG-20121
; REFERENCE/DOCKET NUMBER: USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-433-4150
; TELEFAX: 415-433-8716
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-692C-4

Query Match 98.7%; Score 1413; DB 1; Length 289;
Best Local Similarity 98.6%; Pred. No. 3.2e-137;
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLIILTLFTLTPAVEGDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRFLVLSLLIILTLFTLTPAVEGDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIIMGYRAGDTSYFFNGASATEAAKYVK 120
DB 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIIMGYRAGDTSYFFNEASATEAAKYVK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
QY 181 TSEAAKYKFIQQIGKRVDTKFLPSLAIIISLNSWSALSQIQIASTNNGQFESPVLIN 240
DB 181 TSEAAKYKFIQQIGKRVDTKFLPSLAIIISLNSWSALSQIQIASTNNGQFETPVWLIN 240
QY 241 AQNRVMTITNDAGVVTNSIALLLNRNNMAAMDDVPMTQSFSGCGSYAL 289
DB 241 AQNRVMTITNDAGVVTNSIALLLNRNNMAAMDDVPMTQSFSGCGSYAI 289

RESULT 2
US-08-184-237-4
; Sequence 4, Application US/08184237
; Patent No. 5589367
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantham, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann Myers
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CAL
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,237
; FILING DATE:
; CLASSIFICATION: 435
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 923,692
; FILING DATE: 31-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 600,244
; FILING DATE: 22-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 641,617
; FILING DATE: 16-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 310,881
; FILING DATE: 17-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 160,766
; FILING DATE: 26-FEB-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 160,771
; FILING DATE: 26-FEB-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 347,637
; FILING DATE: 05-MAY-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 363,138
; FILING DATE: 08-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 219,279
; FILING DATE: 15-JUL-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: BIOG-20121 USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-433-4150
; TELEFAX: 415-433-8716
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-184-237-4

Query Match 98.7%; Score 1413; DB 1; Length 289;
Best Local Similarity 98.6%; Pred. No. 3.2e-137;
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLIILTLFTLTPAVEGDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRFLVLSLLIILTLFTLTPAVEGDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIIMGYRAGDTSYFFNGASATEAAKYVK 120
DB 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIIMGYRAGDTSYFFNEASATEAAKYVK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
QY 181 TSEAAKYKFIQQIGKRVDTKFLPSLAIIISLNSWSALSQIQIASTNNGQFESPVLIN 240
DB 181 TSEAAKYKFIQQIGKRVDTKFLPSLAIIISLNSWSALSQIQIASTNNGQFETPVWLIN 240
QY 241 AQNRVMTITNDAGVVTNSIALLLNRNNMAAMDDVPMTQSFSGCGSYAL 289
DB 241 AQNRVMTITNDAGVVTNSIALLLNRNNMAAMDDVPMTQSFSGCGSYAI 289

RESULT 3
US-08-482-920-4
; Sequence 4, Application US/08482920
; Patent No. 5866785
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
```

APPLICANT: Dawson, William O.
APPLICANT: Grantham, George L.
APPLICANT: Turpen, Thomas H.
APPLICANT: Turpen, Ann Myers
APPLICANT: Garger, Stephen J.
APPLICANT: Grill, Laurence K.
TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 184,237
FILING DATE: 19-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 600,244
FILING DATE: 22-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 641,617
FILING DATE: 16-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 310,881
FILING DATE: 17-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 160,766
FILING DATE: 26-FEB-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 160,771
FILING DATE: 26-FEB-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 347,637
FILING DATE: 05-MAY-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 363,138
FILING DATE: 08-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 219,279
FILING DATE: 15-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 8129-112
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-920-4

Query Match 98.7%; Score 1413; DB 2; Length 289;
Best Local Similarity 98.6%; Pred. No. 3.2e-137;
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MIRFLVLSLLILTLTPAVGDSVFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRFLVLSLLILTLTPAVGDSVFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSSLPGSQRYALHLTNYADETISVAIDVTNVIIMGYRAGDTSYFFNGASATEAAKYVFK 120

Db 61 RSSLPGSQRYALHLTNYADETISVAIDVTNVIIMGYRAGDTSYFFNGASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASALMVLIOQS 180
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASALMVLIOQS 180
QY 181 TSEARYKFEIOQIGKRVDTKFLPSLAIIISLNSWSALSQIOIASTNNGQFETPVVLIN 240
Db 181 TSEARYKFEIOQIGKRVDTKFLPSLAIIISLNSWSALSQIOIASTNNGQFETPVVLIN 240
QY 241 AQNRVTITNDAGVVTNSIALLLNRRNNMAAMDDVPMTQSFQCGSSYAL 289
Db 241 AQNRVTITNDAGVVTNSIALLLNRRNNMAAMDDVPMTQSFQCGSSYAI 289
RESULT 4
US-08-484-341-4
Sequence 4, Application US/08484341
GENERAL INFORMATION:
APPLICANT: Donson, Jon
Dawson, William O.
Grantham, George L.
Turpen, Thomas H.
Turpen, Ann Myers
Garger, Stephen J.
Grill, Laurence K.
TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Limbach & Limbach
STREET: 2001 Ferry Building
CITY: San Francisco
STATE: CAL
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,341
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/184,237
FILING DATE: <Unknown>
APPLICATION NUMBER: US 600,244
FILING DATE: 22-OCT-1990
APPLICATION NUMBER: US 641,617
FILING DATE: 16-JAN-1991
APPLICATION NUMBER: US 310,881
FILING DATE: 17-FEB-1989
APPLICATION NUMBER: US 160,766
FILING DATE: 26-FEB-1988
APPLICATION NUMBER: US 160,771
FILING DATE: 26-FEB-1988
APPLICATION NUMBER: US 347,637
FILING DATE: 05-MAY-1989
APPLICATION NUMBER: US 363,138
FILING DATE: 08-JUN-1989
APPLICATION NUMBER: US 219,279
FILING DATE: 15-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: BIOG-20121 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-433-4150
TELEFAX: 415-433-8716
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-484-341-4

Query Match      98.7%; Score 1413; DB 3; Length 289;
Best Local Similarity 98.6%; Pred. No. 3.2e-137;
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
Db 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

Qy 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVK 120
Db 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVK 120

Qy 121 DAMRKVTLPYSGNVERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQ 180
Db 121 DAMRKVTLPYSGNVERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQ 180

Qy 181 TSEARYKFIEQQIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNGQFETPVVLIN 240
Db 181 TSEARYKFIEQQIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNGQFETPVVLIN 240

Qy 241 AQNRVMTITNDAGVVTNSNIALLLNRNNMAAMDDDDVPMTQSFSGCGSYAL 289
Db 241 AQNRVMTITNDAGVVTNSNIALLLNRNNMAAMDDDDVPMTQSFSGCGSYAI 289

RESULT 5
US-08-483-502-4
; Sequence 4, Application US/08483502
; Patent No. 6284492
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantham, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann M.
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,502
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/739,143
; FILING DATE:
; APPLICATION NUMBER: US 600,244
; FILING DATE: 22-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 641,617
; FILING DATE: 16-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 310,881
; FILING DATE: 17-FEB-1989
; PRIOR APPLICATION DATA:

;
; APPLICATION NUMBER: US 160,766
; FILING DATE: 26-FEB-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 160,771
; FILING DATE: 26-FEB-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 347,637
; FILING DATE: 05-MAY-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 363,138
; FILING DATE: 08-JUN-1989
; APPLICATION DATA:
; APPLICATION NUMBER: US 219,279
; FILING DATE: 15-JUL-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 18604-090574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-483-502-4

Query Match      98.7%; Score 1413; DB 3; Length 289;
Best Local Similarity 98.6%; Pred. No. 3.2e-137;
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
Db 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

Qy 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVK 120
Db 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVK 120

Qy 121 DAMRKVTLPYSGNVERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQ 180
Db 121 DAMRKVTLPYSGNVERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQ 180

Qy 181 TSEARYKFIEQQIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNGQFETPVVLIN 240
Db 181 TSEARYKFIEQQIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNGQFETPVVLIN 240

Qy 241 AQNRVMTITNDAGVVTNSNIALLLNRNNMAAMDDDDVPMTQSFSGCGSYAL 289
Db 241 AQNRVMTITNDAGVVTNSNIALLLNRNNMAAMDDDDVPMTQSFSGCGSYAI 289

RESULT 6
US-09-726-651A-4
; Sequence 4, Application US/09726651A
; Patent No. 6448046
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantham, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann M.
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT ANIMAL VIRUS NUCLEIC ACIDS
; FILE REFERENCE: 008010023CNUS01
; CURRENT APPLICATION NUMBER: US/09/726,651A
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 08/483,502
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/184,237
```

;; PRIORITY FILING DATE: 1994-01-19
;; PRIORITY APPLICATION NUMBER: 07/923,692
;; PRIORITY FILING DATE: 1992-07-31
;; PRIORITY APPLICATION NUMBER: 07/600,244
;; PRIORITY FILING DATE: 1990-10-22
;; PRIORITY APPLICATION NUMBER: 07/641,617
;; PRIORITY FILING DATE: 1991-01-16
;; PRIORITY APPLICATION NUMBER: 07/737,899
;; PRIORITY FILING DATE: 1991-07-26
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO. 4
;; LENGTH: 289
;; TYPE: PRT
;; ORGANISM: Chinese Cucumber alpha-trichosanthin
US-09-726-651A-4

Query Match 98.7%; Score 1413; DB 4; Length 289;
Best Local Similarity 98.6%; Pred. No. 3.2e-137;
Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MRFVLVLSLLILFTLTTPAVEGDVSRFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MRFVLVLSLLILFTLTTPAVEGDVSRFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSLPGSQRYALIHLYNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFK 120
DB 61 RSLPGSQRYALIHLYNVADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
QY 181 TSEARYKFIEQOIGKRVDTKFLPSLAIISLENSWSALSQKIQIASTNNGQFESPVLIN 240
DB 181 TSEARYKFIEQOIGKRVDTKFLPSLAIISLENSWSALSQKIQIASTNNGQFESPVLIN 240
QY 241 AQNRVTITNDAGVVTNSIALLLNRNNMAAMDDDDVPMTQSFSGCSYAL 289
DB 241 AQNRVTITNDAGVVTNSIALLLNRNNMAAMDDDDVPMTQSFSGCSYAI 289

RESULT 7
US-08-378-761A-74
; Sequence 74, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651

;; REFERENCE/DOCKET NUMBER: 38272B
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (317) 337-4846
;; INFORMATION FOR SEQ ID NO: 74:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 267 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-378-761A-74

Query Match 92.2%; Score 1321; DB 1; Length 267;
Best Local Similarity 99.3%; Pred. No. 8.4e-128;
Matches 265; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 23 GDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPGSQRYALIHLYNVADET 82
DB 1 GDVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPGSQRYALIHLYNVADET 60
QY 83 ISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAG 142
DB 61 ISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAG 120
QY 143 KIRENIPGLPALDSAITTLFYNNANSAASALMWLIQSTSEARYKFIEQOIGKRVDTK 202
DB 121 KIRENIPGLPALDSAITTLFYNNANSAASALMWLIQSTSEARYKFIEQOIGKRVDTK 180
QY 203 LPSLAIISLENSWSALSQKIQIASTNNGQFESPVLINNAQNORVTITNDAGVVTNSIAL 262
DB 181 LPSLAIISLENSWSALSQKIQIASTNNGQFESPVLINNAQNORVTITNDAGVVTNSIAL 240
QY 263 LLNRNNMAAMDDDDVPMTQSFSGCSYAL 289
DB 241 LLNRNNMAAMDDDDVPMTQSFSGCSYAI 267

RESULT 8
US-08-485-286-74
; Sequence 74, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B

TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 337-4846
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-286-74

Query Match 92.2%; Score 1321; DB 1; Length 267;
Best Local Similarity 99.3%; Pred. No. 8.4e-128;
Matches 265; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 23 GDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYNADET 82
DB 1 GDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYNADET 60

QY 83 ISVAIDVTNVIIMGYRAGDTSYFFNCGASATEAAKVYFKDAMRKVTLTPYSGNYERLQTAAG 142
DB 61 ISVAIDVTNVIIMGYRAGDTSYFFNCGASATEAAKVYFKDAMRKVTLTPYSGNYERLQTAAG 120

QY 143 KIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKTF 202
DB 121 KIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKTF 180

QY 203 LPSLAISLENSWSALSQIQIQTASTNNGQFESPVLINAQORVTTINVDAGVVTSNIAL 262
DB 181 LPSLAISLENSWSALSQIQIQTASTNNGQFESPVLINAQORVTTINVDAGVVTSNIAL 240

QY 263 LLNRNNMAAMDDVPMTQSFSGGSAI 289
DB 241 LLNRNNMAAMDDVPMTQSFSGGSAI 267

RESULT 9
US-08-488-113B-6
Sequence 6, Application US/08488113B
Patent No. 5744580
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Stephen F.
APPLICANT: Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,113B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/425,336
FILING DATE: 18-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 11022US07/200-70.F3.C2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-113B-6

Query Match 83.9%; Score 1202; DB 1; Length 247;
Best Local Similarity 98.8%; Pred. No. 1.3e-115;
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYNADETI 83
DB 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYNADETI 60

QY 84 SVADVTNVIIMGYRAGDTSYFFNCGASATEAAKVYFKDAMRKVTLTPYSGNYERLQTAAGK 143
DB 61 SVADVTNVIIMGYRAGDTSYFFNCGASATEAAKVYFKDAMRKVTLTPYSGNYERLQTAAGK 120

QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKTF 203
DB 121 IRENIPGLPALDSAITTLFYNNANSAASALMWLIQSTSEAAKYKFIQQIGKRVDTKTF 180

QY 204 PSLAIIISLENSWSALSQIQIQTASTNNGQFESPVLINAQORVTTINVDAGVVTSNIAL 263
DB 181 PSLAIIISLENSWSALSQIQIQTASTNNGQFESPVLINAQORVTTINVDAGVVTSNIAL 240

QY 264 LLNRNNMA 270
DB 241 LLNRNNMA 247

RESULT 10
US-08-477-484B-6
Sequence 6, Application US/08477484B
Patent No. 5756699
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Stephen F.
APPLICANT: Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,484B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/425,336
;; FILING DATE: 18-APR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/064,691
;; FILING DATE: 12-MAY-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/988,430
;; FILING DATE: 09-DEC-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/901,707
;; FILING DATE: 19-JUN-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/787,567
;; FILING DATE: 04-NOV-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McNicholas, Janet M.
;; REGISTRATION NUMBER: 32,918
;; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312/707-8889
;; TELEFAX: 312/707-9155
;; TELEX: 650 388-1248
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 247 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-477-484B-6

Query Match 83.9%; Score 1202; DB 1; Length 247;
Best Local Similarity 98.8%; Pred. No. 1.3e-115;
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPFSQRYALIHLTNYADETI 83
Db |||||
QY 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPFSQRYALIHLTNYADETI 60
Db |||||
QY 84 SVAIDVTNVIYMGYRAGDTSYFFNCGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
Db |||||
QY 61 SVAIDVTNVIYMGYRAGDTSYFFNCGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120
Db |||||
QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFIQQIGKRVDTKFL 203
Db |||||
QY 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFIQQIGKRVDTKFL 180
Db |||||
QY 204 PSLAIISLNSWSALSQKIQIASTNNNGQFSPVVLINAQNRVTITNDAGVVTSNIAL 263
Db |||||
QY 181 PSLAIISLNSWSALSQKIQIASTNNNGQFSPVVLINAQNRVTITNDAGVVTSNIAL 240
Db |||||
QY 264 LNRNMA 270
Db 241 LNRNMA 247

RESULT 11
US-08-646-360-6
; Sequence 6, Application US/08646360
; Patent No. 5837491
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA

;; ZIP: 60661
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/646,360
;; FILING DATE: 13-MAY-1996
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US94/05348
;; FILING DATE: 12-MAY-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/064,691
;; FILING DATE: 12-MAY-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/988,430
;; FILING DATE: 09-DEC-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/901,707
;; FILING DATE: 19-JUN-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/787,567
;; FILING DATE: 04-NOV-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McNicholas, Janet M.
;; REGISTRATION NUMBER: 32,918
;; REFERENCE/DOCKET NUMBER: 200-70.P4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312/707-8889
;; TELEFAX: 312/707-9155
;; TELEX: 650 388-1248
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 247 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-646-360-6

Query Match 83.9%; Score 1202; DB 2; Length 247;
Best Local Similarity 98.8%; Pred. No. 1.3e-115;
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPFSQRYALIHLTNYADETI 83
Db |||||
QY 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPFSQRYALIHLTNYADETI 60
Db |||||
QY 84 SVAIDVTNVIYMGYRAGDTSYFFNCGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
Db |||||
QY 61 SVAIDVTNVIYMGYRAGDTSYFFNCGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120
Db |||||
QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFIQQIGKRVDTKFL 203
Db |||||
QY 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFIQQIGKRVDTKFL 180
Db |||||
QY 204 PSLAIISLNSWSALSQKIQIASTNNNGQFSPVVLINAQNRVTITNDAGVVTSNIAL 263
Db |||||
QY 181 PSLAIISLNSWSALSQKIQIASTNNNGQFSPVVLINAQNRVTITNDAGVVTSNIAL 240
Db |||||
QY 264 LNRNMA 270
Db 241 LNRNMA 247

RESULT 12
US-08-839-765-6
; Sequence 6, Application US/08839765
; Patent No. 6146631
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.

APPLICANT: Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
TITLE OF INVENTION: Proteins
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,765
FILING DATE: 15-APR-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/425,336
FILING DATE: 18-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-839-765-6
Query Match 83.9%; Score 1202; DB 3; Length 247;
Best Local Similarity 98.8%; Pred. No. 1.3e-115;
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALHILTNVADETI 83
DB 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALHILTNVADETI 60
QY 84 SVAIDVTNVIYMGYRAGDTSYFPNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
DB 61 SVAIDVTNVIYMGYRAGDTSYFPNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120
QY 144 IRENIPILGUPALDSATITLFFYNANSAALMWLIQSTSEAAKYKTEQQIGKRVDTFL 203
DB 121 IRENIPILGUPALDSATITLFFYNANSAALMWLIQSTSEAAKYKTEQQIGKRVDTFL 180
QY 204 PSIAIISLNSWALSQIQIASTNNGQFSPVVLINAEQNRVITINVDAGVVTNSIAL 263
DB 181 PSIAIISLNSWALSQIQIASTNNGQFSPVVLINAEQNRVITINVDAGVVTNSIAL 240
QY 264 LNRNMA 270
|||||

Db 241 LNRNMA 247
RESULT 13
US-09-136-389-6
Sequence 6, Application US/09136389
Patent No. 6146850
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Stephen F.
APPLICANT: Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-136-389-6
Query Match 83.9%; Score 1202; DB 3; Length 247;
Best Local Similarity 98.8%; Pred. No. 1.3e-115;
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALHILTNVADETI 83
DB 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSLPGSQRYALHILTNVADETI 60
QY 84 SVAIDVTNVIYMGYRAGDTSYFPNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
|||||

Db 61 SVAIDVTNVIYMGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120
QY 144 IRENIPGLPALDSAITTLFFYNANSAASALMWLIQSTSEAAKYKFEQQIGKRVDTKFL 203
Db 121 IRENIPGLPALDSAITTLFFYNANSAASALMWLIQSTSEAAKYKFEQQIGKRVDTKFL 180
QY 204 PSIAIISLENSWSALSQIOIASTNNQFESPVLINAQORVTITNVDAGVVTNSIAL 263
Db 181 PSIAIISLENSWSALSQIOIASTNNQFESPVLINAQORVTITNVDAGVVTNSIAL 240
QY 264 LNRNNMA 270
Db 241 LNRNNMA 247

RESULT 14
US-09-610-838-6
; Sequence 6, Application US/09610838
; Patent No. 6376217
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,838
; FILING DATE: 06-JUL-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/136,389
; FILING DATE: 18-AUG-1998
; APPLICATION NUMBER: 08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-610-838-6
Query Match 83.9%; Score 1202; DB 4; Length 247;
Best Local Similarity 98.8%; Pred. No. 1.3e-115;
Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPILRSLPGSQRYALHILTNVADETI 83
Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPILRSLPGSQRYALHILTNVADETI 60
QY 84 SVAIDVTNVIYMGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
Db 61 SVAIDVTNVIYMGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120
QY 144 IRENIPGLPALDSAITTLFFYNANSAASALMWLIQSTSEAAKYKFEQQIGKRVDTKFL 203
Db 121 IRENIPGLPALDSAITTLFFYNANSAASALMWLIQSTSEAAKYKFEQQIGKRVDTKFL 180
QY 204 PSIAIISLENSWSALSQIOIASTNNQFESPVLINAQORVTITNVDAGVVTNSIAL 263
Db 181 PSIAIISLENSWSALSQIOIASTNNQFESPVLINAQORVTITNVDAGVVTNSIAL 240
QY 264 LNRNNMA 270
Db 241 LNRNNMA 247

RESULT 15
US-09-711-485-6
; Sequence 6, Application US/09711485
; Patent No. 6649742
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/711,485
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/839,765
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.

Wed Sep 15 10:32:03 2004

REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 247 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-711-485-6

Query Match 83.9%; Score 1202; DB 4; Length 247;

Best Local Similarity 98.8%; Pred. No. 1.3e-115;

Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 24 DVSRFSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLTNYADETI 83

Db 1 DVSRFSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLTNYADETI 60

QY 84 SVAIDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 143

Db 61 SVAIDVTNVIYIMGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLTPYSGNYERLQTAAGK 120

QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOQSTSEARYKFEIQOIGKRVDTKFL 203

Db 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOQSTSEARYKFEIQOIGKRVDTKFL 180

QY 204 PSIAIISLENSWSALSQKIQIASTNNGQFESPVLINAQNVATITNVDAAGVVTSTNIAL 263

Db 181 PSIAIISLENSWSALSQKIQIASTNNGQFESPVLINAQNVATITNVDAAGVVTSTNIAL 240

QY 264 LNRNNMA 270

Db 241 LNRNNMA 247

Search completed: September 10, 2004, 14:02:34

Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 13:24:40 ; Search time 121 Seconds
(without alignments)
674.845 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MRFLVLSLLILTLFTTPA.....AAMDDVPMTQSGGCGSYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A: Geneseq 26Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1424	99.4	289	5	ABB07660 Native tr
2	1422	99.3	289	4	AAB99329 Trichosan
3	1413	98.7	289	2	AAR07514 Trichosan
4	1413	98.7	289	2	AAR25572 Trichosan
5	1413	98.7	289	2	AAR29272 Trichosan
6	1413	98.7	289	2	AAR55129 Alpha-Tri
7	1413	98.7	289	2	AAR10468 Chinese c
8	1413	98.7	289	2	AAR11870 Chinese c
9	1413	98.7	289	2	AAR01374 Chinese c
10	1413	98.7	289	2	AAR84192 Chinese c
11	1413	98.7	289	3	AAR87791 Chinese c
12	1410	98.5	289	2	AAR32986 Encodes c
13	1326	92.6	289	2	AAR07523 Alpha-Tri
14	1326	92.6	289	2	AAR29276 Ribosome
15	1321	92.2	267	2	AAR21703 Trichosan
16	1321	92.2	267	2	AAR25140 Trichosan
17	1214	84.8	247	3	AAY69048 Amino aci
18	1214	84.8	248	2	AAR07518 Synthetic
19	1214	84.8	248	2	AAR25573 Mature al
20	1211	84.6	247	2	AAR67359 Trichosan
21	1183	82.6	247	2	AAR21605 Antitumou
22	1164.5	81.3	290	2	AAR92481 Bryodin 1
23	1144	79.9	255	2	AAR74181 Type I ri
24	1139	79.5	246	2	AAR52636 Bioactive
25	1130	78.9	255	2	AAR37295 Plant typ

26	1130	78.9	255	2	AAR63907 Type I ri
27	1089.5	76.1	287	3	AAB01299 Wild type
28	1074	75.0	496	4	AAB36828 BDI-G28.5
29	1064	74.3	247	3	AAY69043 Amino aci
30	1064	74.3	248	4	AAB36824 Residues
31	1062.5	74.2	267	3	AAB01300 Altered b
32	806.5	56.3	263	2	AAR63908 Type I ri
33	806.5	56.3	263	2	AAR74182 Type I ri
34	802.5	56.0	277	2	AAR29909 Prod. of
35	799.5	55.8	283	2	AAR37296 Plant typ
36	798	55.7	280	2	AAR07520 Alpha-Tri
37	798	55.7	280	2	AAR07521 Alpha-Tri
38	798	55.7	280	2	AAR25575 Ribosome
39	798	55.7	280	2	AAR25576 Ribosome
40	786.5	54.9	272	2	AAR07522 Alpha-Tri
41	786.5	54.9	272	2	AAR25577 Ribosome
42	775	54.1	278	2	AAR29910 Prod. of
43	772.5	53.9	276	2	AAR53731 Luffin (r
44	760	53.1	250	2	AAR21707 Momordin.
45	760	53.1	250	2	AAR25144 Momordin

ALIGNMENTS

RESULT 1
ABB07660
ID ABB07660 standard; protein; 289 AA.
XX
AC ABB07660;
XX
DT 20-MAY-2002 (first entry)
XX
DE Native trichosanthin (TCS) protein sequence.
XX

Trichosanthin; TCS; mutant; MTCS; therapeutic; cytostatic; antitumour;
anti-human immunodeficiency virus; virucide; immunostimulant; Gene;
ectopic pregnancy.

Trichosanthin kirilowii.

Key	Location/Qualifiers
Peptide	1..23
Misc-difference	/note= "signal peptide"
Protein	24..289
Region	/note= "mature protein"
Misc-difference	174..180
Region	/note= "MTCS contains a modification of at least one amino acid residue in this region"
Misc-difference	203..226
Region	/note= "encoded by ATT"
Misc-difference	230..244
Region	/note= "MTCS contains a modification of at least one amino acid residue in this region"
Misc-difference	289
Region	/note= "encode dby ATT"

W0200212537-A2.

14-FEB-2002.

18-JUL-2001; 2001WO-CN001178.

02-AUG-2000; 2000CN-00119553.

18-JAN-2001; 2001CN-00103102.

(BEIJ-) BEIJING STM BIOTECH LTD.
(KEYJ/) KE Y.

14-FEB-2001.
02-AUG-2000; 2000CN-00119553.
02-AUG-2000; 2000CN-00119553.
(SHAN-) SHANGHAI INST CYTOBIOLOGY CHINESE ACAD.
Ke Y, Nie H;
WPI; 2001-291745/31.
N-PSDB; AAH41473.
Trichosanthin mutant and its preparing process.
Disclosure; Page 2 (disclosure); 15pp; Chinese.
The present invention describes a trichosanthin mutant which is prepared through the mutational deformation of the trichosanthin gene and using an expression system. The trichosanthin gene is isolated from Trichosanthes kirilowii (Mongolian snake-gourd). The trichosanthin mutant has several advantages including high bioactivity, high selectivity to target, and strong kill action to cancer cells, virus and HIV. It can also be used for metaphase induced labour. The present sequence represents the protein sequence of wild type trichosanthin which is given in the exemplification of the present invention
Query Match 99.3%; Score 1422; DB 4; Length 289;
Best Local Similarity 99.3%; Pred. No. 2.6e-127;
Matches 287; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSSLPGSORYALIHLTNYADETISVAIDVTNVYINGYRAGDTSYFFNGASATEAAKYVFK 120
DB 61 RSSLPGSORYALIHLTNYADETISVAIDVTNVYINGYRAGDTSYFFNGASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
QY 181 TSEAAARYKFEQIQKRVDTKFLPSLAIISLNSWSALSQIQIASTNNQFESPVLIN 240
DB 181 TSEAAARYKFEQIQKRVDTKFLPSLAIISLNSWSALSQIQIASTNNQFESPVLIN 240
QY 241 AQNRVTITNDAGVVTNSIALLLNRRNMAAMDDVPMTQSFSGCGSYAL 289
DB 241 AQNRVTITNDAGVVTNSIALLLNRRNMAAMDDVPMTQSFSGCGSYAI 289
RESULT 3
AAR07514
ID AAR07514 standard; protein; 289 AA.
XX AAR07514;
AC AAR07514;
XX 06-FEB-1991 (first entry)
DT 06-FEB-1991 (first entry)
XX Trichosanthin from Trichosanthes kirilowii.
DE trichosanthin (TCS); Human Immunodeficiency Virus; inhibition.
XX Trichosanthes kirilowii.
XX Key Location/Qualifiers
FH 1. .21
FT /label= signal peptide
FT /note= "hydrophobic"
XX

14-FEB-2001.
02-AUG-2000; 2000CN-00119553.
02-AUG-2000; 2000CN-00119553.
(SHAN-) SHANGHAI INST CYTOBIOLOGY CHINESE ACAD.
Ke Y, Nie H;
WPI; 2001-291745/31.
N-PSDB; AAH41473.
Trichosanthin mutant and its preparing process.
Disclosure; Page 2 (disclosure); 15pp; Chinese.
The present invention describes a trichosanthin mutant which is prepared through the mutational deformation of the trichosanthin gene and using an expression system. The trichosanthin gene is isolated from Trichosanthes kirilowii (Mongolian snake-gourd). The trichosanthin mutant has several advantages including high bioactivity, high selectivity to target, and strong kill action to cancer cells, virus and HIV. It can also be used for metaphase induced labour. The present sequence represents the protein sequence of wild type trichosanthin which is given in the exemplification of the present invention
Query Match 99.4%; Score 1424; DB 5; Length 289;
Best Local Similarity 99.7%; Pred. No. 1.7e-127;
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRELVLSLLILTLFTTPAVEGDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSSLPGSORYALIHLTNYADETISVAIDVTNVYINGYRAGDTSYFFNGASATEAAKYVFK 120
DB 61 RSSLPGSORYALIHLTNYADETISVAIDVTNVYINGYRAGDTSYFFNGASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLIQS 180
QY 181 TSEAAARYKFEQIQKRVDTKFLPSLAIISLNSWSALSQIQIASTNNQFESPVLIN 240
DB 181 TSEAAARYKFEQIQKRVDTKFLPSLAIISLNSWSALSQIQIASTNNQFESPVLIN 240
QY 241 AQNRVTITNDAGVVTNSIALLLNRRNMAAMDDVPMTQSFSGCGSYAL 289
DB 241 AQNRVTITNDAGVVTNSIALLLNRRNMAAMDDVPMTQSFSGCGSYAL 289
RESULT 2
AAB99329
ID AAB99329 standard; protein; 289 AA.
XX AAB99329;
AC AAB99329;
XX 23-AUG-2001 (first entry)
DT 23-AUG-2001 (first entry)
XX Trichosanthes kirilowii trichosanthin (TCS) protein sequence.
DE Trichosanthes kirilowii; trichosanthin; TCS; mutagenesis; mutation; MTCS;
KW mutant of trichosanthin; Mongolian snake-gourd; bioactivity; selectivity;
KW cancer; virus; HIV; metaphase induced labour.
XX Trichosanthes kirilowii.
OS Trichosanthes kirilowii.
XX Key Location/Qualifiers
FH 230
FT Misc-difference
FT /note= "encoded by GCA"
XX CN1283630-A.
XX

PN WO9012097-A.
 XX 18-OCT-1990.
 XX 04-APR-1989; 89US-003333184.
 XX 04-APR-1989; 89US-003333184.
 XX (GENE-) GENELABS INC.
 XX Platek M, Chow T, Fry K;
 XX WPI; 1990-334847/44.
 DR N-PSDB; AAQ06343.
 XX Recombinant tri:chosanthin protein - with selective inhibitory effect on
 FT viral expression in HIV infected T-cells or monocyte-macrophase.
 XX Example; Fig 4; 102pp; English.
 CC Genomic DNA was isolated from T.kirilowii leaves from Korea and a library
 CC was constructed. Clone pQ21D was identified as likely to contain a TCS-
 CC encoding sequence in its 4kb insert. The deduced amino acid sequence is
 CC identical to that of TCS purified from Cantonese T.kirilowii roots,
 CC except for 2 conservative substitu- tions, i.e. Thr for Ser at position
 CC 211 and Met for Thr at position 224. The Canton protein lacks the last 19
 CC C-terminal amino acid residues. See also AAQ06344-Q06351
 XX Sequence 289 AA;
 SQ

Query Match 98.7%; Score 1413; DB 2; Length 289;
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLTPAVEGDVSRFLSGATSSSYGVFISNLKALPNERKLYDIPLL 60
 Db 1 MIRFLVLSLLILTLTPAVEGDVSRFLSGATSSSYGVFISNLKALPNERKLYDIPLL 60
 QY 61 RSLPGSQRYALIHNTVADETISVAIDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVFK 120
 Db 61 RSLPGSQRYALIHNTVADETISVAIDVTNVIYIMGYRAGDTSYFFNGASATEAAKYVFK 120
 QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENTPLGLPALDSAITTLFYNNANSAASALMVLIIQS 180
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENTPLGLPALDSAITTLFYNNANSAASALMVLIIQS 180
 QY 181 TSEAAKYKFIQQIGKRVDTKFLPSLAIISLNSWSLSKQIQIASTNNGOFESPVLIN 240
 Db 181 TSEAAKYKFIQQIGKRVDTKFLPSLAIISLNSWSLSKQIQIASTNNGOFESPVLIN 240
 QY 241 AQONQVITNVDAGVTSNIALLNRRNNMAAMDDVPMTQSGGGSVAL 289
 Db 241 AQONQVITNVDAGVTSNIALLNRRNNMAAMDDVPMTQSGGGSVAL 289

RESULT 4
 AAR25572
 ID AAR25572 standard; protein; 289 AA.
 XX
 AC AAR25572;
 XX
 DT 25-MAR-2003 (revised)
 DT 13-JAN-1993 (first entry)
 XX
 DE Trichosanthin from Trichosanthin kirilowii.
 XX
 KW TCS; alpha-trichosanthin; Radix trichosanthin; abortifacient;
 KW ribosome inactivating protein; RIP; HIV-infected human T cells;
 KW human immunodeficiency virus.
 XX
 OS Trichosanthin kirilowii.
 XX
 FH Key Location/Qualifiers

Misc-difference 57 /note= "Leu in previously published sequence"
 FT
 FT Misc-difference 60 /note= "Ile in previously published sequence"
 FT
 FT Misc-difference 72 /note= "Ile in previously published sequence"
 FT
 FT Misc-difference 82..84 /note= "Ile in previously published sequence"
 FT
 FT Misc-difference 82..84 /note= "region not present in previously published
 FT sequence"
 FT
 FT Misc-difference 92..93 /note= "previously published sequence contained a 10
 FT amino acid insert (DAGLPNRAVL) between Val and Tyr"
 FT
 FT Misc-difference 143 /note= "Gly in previously published sequence"
 FT
 FT Misc-difference 144 /note= "Leu in previously published sequence"
 FT
 FT Misc-difference 196 /note= "Ser in previously published sequence"
 FT
 FT Misc-difference 214..215 /note= "previously published sequence contained a Leu
 FT inserted between Ser and Trp"
 FT
 FT Misc-difference 216 /note= "Leu in previously published sequence"
 FT
 FT Misc-difference 231 /note= "Thr in previously published sequence"
 FT
 FT Misc-difference 234 /note= "Ser in directly sequenced TCS"
 FT
 FT Misc-difference 246..266 /note= "21 amino acids not present in previously
 FT published sequence"
 FT
 FT Misc-difference 247 /note= "Thr in directly sequenced TCS"
 FT
 XX US5128460-A.
 XX
 XX 07-JUL-1992.
 XX
 PF 04-APR-1990; 90US-00504775.
 XX
 PR 04-APR-1989; 89US-003333184.
 PR 07-SEP-1989; 89US-00404326.
 XX
 XX (GENE-) GENELABS INC.
 XX
 XX Piatak M, Chow TP, Fry K;
 XX
 DR WPI; 1992-249485/30.
 DR N-PSDB; AAQ26499.
 XX
 XX Nucleic acid encoding trichosanthin protein - which can be used to
 FT inactivate ribosome(s) to inhibit protein synthesis or to inhibit HIV
 FT expression.
 XX
 PS Claim 1; Fig 4; 53pp; English.
 XX
 CC Trichosanthin protein was isolated from T.kirilowii root tuber and
 CC sequenced. The amino acid sequence was used to design sets of degenerate
 CC primers (see AAQ26506-8) for PCR amplification of the TCS coding
 CC sequence. The amplified product was used as a probe to isolate TCS coding
 CC sequence from T.kirilowii genomic libraries. One clone (pQ21D) contained
 CC a 4kb insert. The amino acid sequence deduced from the pQ21D insert
 CC differed from the purified TCS sequence by conservative substitutions at
 CC two positions (see Features Table). The differences are postulated to
 CC indicate minor variations between strains; the purified TCS was obtained
 CC from the Canton region of China and the genomic DNA was obtained from
 CC T.kirilowii leaves from Korea. The amino acid sequence was also found to
 CC differ substantially from the previously published TCS sequence (Acta
 CC Chemica Sinica, 43:1943, 1984 and Pure and Appl. Chem., 58(5):789, 1986).
 CC The differences are shown in the Features Table. The present sequence
 CC agrees closely with X-ray diffraction data on crystallized TCS. The 21-
 CC amino acid insert also provides greater sequence homology with a number
 CC of RIPS such as ricin A chain and abrin A chain than the previously
 CC published sequence. (Updated on 25-MAR-2003 to correct PF field.)

CC (Updated on 25-MAR-2003 to correct PR field.)
 XX Sequence 289 AA;
 SQ

Query Match 98.7%; Score 1413; DB 2; Length 289;
 Best Local Similarity 98.6%; Pred. No. 1.9e-126; Indels 0; Gaps 0;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MIRFLVLSLLILFLTPPAVEGDVSRFGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 Db 1 MIRFLVLSLLILFLTPPAVEGDVSRFGATSSSYGVFISNLRKALPNERKLYDIPLL 60

Qy 61 RSSLPQSQRVALIHLTNVADETISVAIDVTNVIMGYRAGDTSYFFNGASATEAKYVFK 120
 Db 61 RSSLPQSQRVALIHLTNVADETISVAIDVTNVIMGYRAGDTSYFFNEASATEAKYVFK 120

Qy 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQS 180
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQS 180

Qy 181 TSEAAARYKFIEQQIGKRVDTFLPSLAIISLENSWSALSCKOIQIASTNNGQFETPVVLIN 240
 Db 181 TSEAAARYKFIEQQIGKRVDTFLPSLAIISLENSWSALSCKOIQIASTNNGQFETPVVLIN 240

Qy 241 AQONRVITITNDAGVVTNSIALLLNNNNMAAMDDVPMTQSGCGSYAL 289
 Db 241 AQONRVITITNDAGVVTNSIALLLNNNNMAAMDDVPMTQSGCGSYAI 289

RESULT 5
 AAR29272
 ID AAR29272 standard; protein; 289 AA.
 XX
 AC AAR29272;
 XX
 FT Key Location/Qualifiers
 FT Peptide 1..23
 FT /label= sig_peptide
 FT /note= "putative N-terminal extension of the mature TCS"
 FT Protein 24..270
 FT /label= mat_protein
 FT Misc-difference 234
 FT /note= "plant-derived TCS has Ser at this position"
 FT Misc-difference 247
 FT /note= "plant-derived TCS has Thr at this position"
 FT Protein 270..289
 FT /note= "putative C-terminal extension of the mature TCS"
 FT
 XX US5166056-A.
 XX
 PD 24-NOV-1992.
 XX
 XX 09-DEC-1991; 91US-00804293.
 XX
 XX 04-APR-1989; 89US-00333184.
 PR 07-SEP-1989; 89US-00404326.
 XX
 XX (GENE-) GENELABS INC.
 XX
 XX Piatak M, Chow TP;
 XX
 XX WPI; 1992-414954/50.
 DR
 XX

PT Recombinant Trichosanthin protein prodn. in E. coli - for use in the
 PT selective inhibition of viral expression in HIV infected cells.
 XX
 PS Disclosure; Fig 4; 37pp; English.
 XX

The sequence is identical to that of plant-derived TCS except for two
 conservative changes: a Thr for a Ser substitution at position 211 and a
 Met for a Thr substitution at position 224. TCS is likely produced as a
 secreted protein that undergoes post-translational processing at both the
 amino and carboxy ends. The TCS coding sequence was amplified using the
 primers of AAQ31828-30. The amplified prod. has the sequence of AAQ31827,
 which was used as a probe. One isolate, pQ21D, comprises the sequence of
 AAQ31826. The recombinant TCS sequence may be used in the recombinant
 prodn. of TCS. TCS can be used for the selective inhibition of viral
 expression in HIV-infected human T-cells or macrophages. (Updated on 25-
 MAR-2003 to correct PF field.)

XX Sequence 289 AA;
 SQ

Query Match 98.7%; Score 1413; DB 2; Length 289;
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MIRFLVLSLLILFLTPPAVEGDVSRFGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 Db 1 MIRFLVLSLLILFLTPPAVEGDVSRFGATSSSYGVFISNLRKALPNERKLYDIPLL 60

Qy 61 RSSLPQSQRVALIHLTNVADETISVAIDVTNVIMGYRAGDTSYFFNGASATEAKYVFK 120
 Db 61 RSSLPQSQRVALIHLTNVADETISVAIDVTNVIMGYRAGDTSYFFNEASATEAKYVFK 120

Qy 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQS 180
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMWLIQS 180

Qy 181 TSEAAARYKFIEQQIGKRVDTFLPSLAIISLENSWSALSCKOIQIASTNNGQFETPVVLIN 240
 Db 181 TSEAAARYKFIEQQIGKRVDTFLPSLAIISLENSWSALSCKOIQIASTNNGQFETPVVLIN 240

Qy 241 AQONRVITITNDAGVVTNSIALLLNNNNMAAMDDVPMTQSGCGSYAL 289
 Db 241 AQONRVITITNDAGVVTNSIALLLNNNNMAAMDDVPMTQSGCGSYAI 289

RESULT 6
 AAR55129
 ID AAR55129 standard; protein; 289 AA.
 XX
 AC AAR55129;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 11-JAN-1995 (first entry)
 XX
 XX Alpha-trichosanthin coding.
 XX
 XX Virus; recombination; plant virus; alpha trichosanthin; phenotype;
 KW alpha amylase; alpha haemoglobin; brome mosaic virus; gemini virus;
 KW rice necrosis virus tobamovirus; gene expression; chinese cucumber.
 XX
 XX Trichosanthes kirilowii.
 XX
 XX US5316931-A.
 XX
 XX 31-MAY-1994.
 XX
 XX 31-JUL-1992; 92US-00923692.
 XX
 XX 26-FEB-1988; 88US-00160766.
 PR 26-FEB-1988; 88US-00160771.
 PR 15-JUL-1988; 89US-00219279.
 PR 17-FEB-1989; 89US-00310881.
 PR 05-MAY-1989; 89US-00347637.

PR 08-JUN-1989; 89US-00363138.
 PR 22-OCT-1990; 90US-00600244.
 PR 16-JAN-1991; 91US-00641617.
 PR 26-JUL-1991; 91US-00737899.
 PR 01-AUG-1991; 91US-00739143.
 XX
 PA (BIOS-) BIOSOURCE GENETICS CORP.
 XX
 PI Garger SJ, Turpen AM, Grill LK, Grantham GL, Dawson WO, Donson J;
 PI Turpen TH;
 XX
 DR WPI; 1994-176269/21.
 DR N-PSDB; AAQ65573.
 XX
 PT New recombinant plant viral nucleic acid - capable of systemic infection
 PT and stable expression of non-native nucleic acid in plant host.
 XX
 PS Example 4; Col 47-50; 44pp; English.
 XX
 CC The alpha-trichosanthin gene may be inserted into a recombinant plant
 CC virus which can then be used to infect plants for the production of non-
 CC native products (in this case alpha-trichosanthin). Other genes which
 CC may be inserted into the virus are those which control a phenotypic
 CC trait, such as male sterility, or sequences encoding anti-sense RNA which
 CC can be useful to prevent the expression of undesired phenotypic traits.
 CC The recombinant virus is derived from a plus sense, single stranded virus
 CC selected from tobamovirus, bromo mosaic virus, rice necrosis virus or a
 CC gemini virus. (Updated on 25-MAR-2003 to correct PF field.) (Updated on
 CC 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 289 AA;
 Query Match 98.7%; Score 1413; DB 2; Length 289;
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MIRFLVLSLLILTLFTTPAVGDSVFRISGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 Db 1 MIRFLVLSLLILTLFTTPAVGDSVFRISGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 QY 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120
 Db 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120
 QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
 QY 181 TSEAAKYKFIQQIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240
 Db 181 TSEAAKYKFIQQIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240
 QY 241 AQORVITITNDAGVTSNIALLNRRNNMAAMDDVPMTQSFSGCSYAL 289
 Db 241 AQORVITITNDAGVTSNIALLNRRNNMAAMDDVPMTQSFSGCSYAI 289
 RESULT 7
 AAW10468
 ID AAW10468 standard; protein; 289 AA.
 XX
 AC AAW10468;
 XX
 DT 17-OCT-2003 (revised)
 DT 26-APR-1997 (first entry)
 XX
 DE Chinese cucumber alpha-trichosanthin.
 XX
 KW Recombinant viral nucleic acid; RNA virus; vector; tobacco mosaic virus;
 KW TWV; Chinese cucumber; alpha-trichosanthin;
 KW ribosome inactivating protein; antiviral; virucide; transgenic plant.
 XX
 OS Trichosanthin kirilowii; Maximowicz.

XX WO9640867-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 06-JUN-1996; 96WO-US009299.
 XX
 PR 07-JUN-1995; 95US-00483502.
 XX
 PA (BIOS-) BIOSOURCE TECHNOLOGIES INC.
 XX
 PI Donson J, Dawson WO, Grantham GL, Turpen TH, Turpen AM;
 PI Garger SJ, Grill LK;
 XX
 DR WPI; 1997-065181/06.
 DR N-PSDB; AAT47094.
 XX
 PT Recombinant viral nucleic acid producing, e.g. male sterility in plants -
 PT comprises nucleic acid whose transcription is controlled by another
 PT sequence.
 XX
 PS Example 4; Page 124-125; 149pp; English.
 XX
 CC Chinese cucumber alpha-trichosanthin (AAM10468) is a ribosome
 CC inactivating protein of potential use in the treatment of HIV infection.
 CC Expression vector pBGL152 was constructed in which the alpha-
 CC trichosanthin coding sequence (see also AAT47094) was placed under
 CC control of the promoter of the tobacco mosaic virus-U1 coat protein gene,
 CC which had been deleted. The viral nucleic acid was capable of self-
 CC replication, encapsidation and systemic spread in infected Nicotiana
 CC benthamiana plants, and directed the high-level expression of
 CC biologically active alpha-trichosanthin in plant tissues. (Updated on 17-
 CC OCT-2003 to standardise OS field)
 XX
 SQ Sequence 289 AA;
 Query Match 98.7%; Score 1413; DB 2; Length 289;
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MIRFLVLSLLILTLFTTPAVGDSVFRISGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 Db 1 MIRFLVLSLLILTLFTTPAVGDSVFRISGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 QY 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120
 Db 61 RSSLPGSQRYALIHLTNYADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120
 QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180
 QY 181 TSEAAKYKFIQQIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240
 Db 181 TSEAAKYKFIQQIGKRVKDTFLPSLAIISLNSWSALSQIQIASTNNGQFESPVLIN 240
 QY 241 AQORVITITNDAGVTSNIALLNRRNNMAAMDDVPMTQSFSGCSYAL 289
 Db 241 AQORVITITNDAGVTSNIALLNRRNNMAAMDDVPMTQSFSGCSYAI 289
 RESULT 8
 AAW11870
 ID AAW11870 standard; protein; 289 AA.
 XX
 AC AAW11870;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 21-APR-1997 (first entry)
 XX
 DE Chinese cucumber alpha-trichosanthin.
 XX

KW	Recombinant virus; alpha-haemoglobin; human; chinese cucumber;	181	TSEARYKEIEQIGKRVKDTFLPSLAIIISLNSALSQIQIASTNNQFETPVVLIN	240
KW	alpha-trichosanthin; rice; alpha amylase; beta-haemoglobin;	241	ACQQRVTITNDAGVVTNSIALLLNNNMAAMDDVPMTQSCGGSYAL	289
KW	subgenomic promoter; coat protein.	241	ACQQRVTITNDAGVVTNSIALLLNNNMAAMDDVPMTQSCGGSYAI	289
OS	Cucumis sp.			
XX	US5589367-A.			
XX	31-DEC-1996.			
XX	19-JAN-1994;	94US-00184237.		
XX	26-FEB-1988;	88US-00160766.		
XX	26-FEB-1988;	88US-00160771.		
XX	15-JUL-1988;	88US-00219279.		
XX	17-FEB-1989;	89US-00310881.		
XX	05-MAY-1989;	89US-00347637.		
XX	08-JUN-1989;	89US-00363138.		
XX	22-OCT-1990;	90US-00600244.		
XX	16-JAN-1991;	91US-00641617.		
XX	26-JUL-1991;	91US-00737899.		
XX	01-AUG-1991;	91US-00739143.		
XX	31-JUL-1992;	92US-00923692.		
XX	(BIOS-) BIOSOURCE TECHNOLOGIES INC.			
XX	Garger SJ, Turpen AM, Grill LK, Grantham GL, Donson J, Dawson WO;			
XX	Turpen TH;			
XX	WPI; 1997-076845/07.			
XX	N-PSDB; AAT61376.			
XX	Recombinant viral DNA for altering plant phenotype or protein prodn -			
XX	contains non-native sub-genomic promoter for expression of heterologous			
XX	protein and native promoter for expression of coat protein.			
XX	Example 4; Col 45-46; 42pp; English.			
XX	The sequences given in AAW11868-71 represent proteins which were produced			
XX	by the recombinant viruses of the invention. The viruses are recombinant			
XX	plant viruses which comprise a native plant virus subgenomic promoter, at			
XX	least one non-native plant virus subgenomic promoter, and a sequence			
XX	encoding a plant virus coat protein. These heterologous sequences are			
XX	preferably under the control of the native promoter sequence. By using a			
XX	plant virus existing cells can be altered with a new coding sequences			
XX	without involving germ cell. The recombinant viruses are stable and can			
XX	cause systemic infection, with stable expression/transcription in plants			
XX	that are hosts for the non-native part of the vector. The nucleotide			
XX	sequences encoding these protein preferably integrated in plant viruses			
XX	having either the O-coat protein or the UI-coat protein gene. (Updated on			
XX	25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR			
XX	field.) (Updated on 27-AUG-2003 to correct OS field.)			
XX	Sequence 289 AA;			
XX	Query Match	98.7%;	Score 1413;	DB 2; Length 289;
XX	Best Local Similarity	98.6%;	Pred. No. 1.9e-126;	
XX	Matches 285; Conservative	2;	Mismatches	2; Indels 0; Gaps 0;
QY	1	MIRFLVLSLLILFLTTTTPAVEGVDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL	60	
DB	1	MIRFLVLSLLILFLTTTTPAVEGVDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL	60	
QY	61	RSLPQSQRVALIHLTNVADETTSVADITNVIMYGRAGTSSYFENGASATAAKVFK	120	
DB	61	RSLPQSQRVALIHLTNVADETTSVADITNVIMYGRAGTSSYFENGASATAAKVFK	120	
QY	121	DAMRKVTLPSYGNVRLQTAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS	180	
DB	121	DAMRKVTLPSYGNVRLQTAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS	180	
QY	181	TSEARYKEIEQIGKRVKDTFLPSLAIIISLNSALSQIQIASTNNQFETPVVLIN	240	

XX SQ Sequence 289 AA;

Query Match 98.7%; Score 1413; DB 2; Length 289;
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 Db 1 MIRFLVLSLLILTLFTTPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPQSQRVALIHLTNVADETISVAIDVTNNVIMGYRAGDTSYFFNGASATEAAKYVFK 120
 Db 61 RSSLPQSQRVALIHLTNVADETISVAIDVTNNVIMGYRAGDTSYFFNGASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANGASALMWLIQS 180
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANGASALMWLIQS 180

QY 181 TSEAAKYKFIQQIGKRVDTFLPSLAIIISLNSWSALSQIQIASTNNQFETPVVLIN 240
 Db 181 TSEAAKYKFIQQIGKRVDTFLPSLAIIISLNSWSALSQIQIASTNNQFETPVVLIN 240

QY 241 AQNRVLTITNDAGVVTNSIALLLNRNNMAAMDDDDVPMTQSGCGSYAL 289
 Db 241 AQNRVLTITNDAGVVTNSIALLLNRNNMAAMDDDDVPMTQSGCGSYAI 289

RESULT 10
 AAW84192
 ID AAW84192 standard; protein; 289 AA.

XX AC AAW84192;
 XX 27-AUG-2003 (revised)
 DT 01-APR-1999 (first entry)
 XX Chinese cucumber alpha-trichosanthin.

KW Chinese cucumber; alpha-trichosanthin; plant virus;
 KW RNA plant virus promoter; systemic infection; foreign gene expression;
 KW AIDS therapeutic drug.

XX Momordica cochinchinensis.
 XX US5866785-A.
 XX 02-FEB-1999.

XX 07-JUN-1995; 95US-00482920.
 XX 26-FEB-1988; 88US-00160766.
 XX 26-FEB-1988; 88US-00160771.
 XX 15-JUL-1988; 88US-00219279.
 XX 17-FEB-1989; 89US-00310881.
 XX 05-MAY-1989; 89US-00347637.
 XX 08-JUN-1989; 89US-00363138.
 XX 22-OCT-1990; 90US-00600244.
 XX 16-JAN-1991; 91US-00641617.
 XX 26-JUL-1991; 91US-00737899.
 XX 01-AUG-1991; 91US-00739143.
 XX 31-JUL-1992; 92US-00923692.
 XX 19-JAN-1994; 94US-00184237.

XX (BIOS-) BIOSOURCE TECHNOLOGIES INC.
 XX Garger SJ, Turpen AM, Grill LK, Grantham GL, Donson J, Dawson WO;
 XX Turpen TH;
 XX WPI; 1999-142035/12.
 XX N-PSDB; AAX03385.

XX Recombinant plant viral vector - that is capable of systemic infection in

PT host plant and stable production of heterologous DNA useful for producing
 PT therapeutic proteins for treating e.g. AIDS.
 PS Example 4; Col 45-48; 45pp; English.
 XX The present sequence represents chinese cucumber alpha-trichosanthin. The
 CC nucleic acid sequence can be expressed in the plant viral constructs of
 CC the invention. The specification describes a recombinant plant viral
 CC nucleic acid derived from a positive (+)-sense RNA plant virus comprising
 CC a native (+)-sense RNA plant virus promoter that is linked to an
 CC expression sequence and a heterologous (+)-sense RNA plant virus promoter
 CC that is linked to an expression sequence. The promoters are incapable of
 CC recombination with each other, and one of the expression sequences
 CC encodes a plant viral coat protein while the other is optionally a
 CC heterologous coding sequence. The plant viral nucleic acid is capable of
 CC systemic infection in a host plant. The viral construct is useful for the
 CC introduction and expression of non-viral foreign genes in plants and the
 CC production of e.g. potential AIDS therapeutic drugs. (Updated on 27-AUG-
 CC 2003 to correct Os field.)
 XX Sequence 289 AA;
 SQ

Query Match 98.7%; Score 1413; DB 2; Length 289;
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 Db 1 MIRFLVLSLLILTLFTTPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPQSQRVALIHLTNVADETISVAIDVTNNVIMGYRAGDTSYFFNGASATEAAKYVFK 120
 Db 61 RSSLPQSQRVALIHLTNVADETISVAIDVTNNVIMGYRAGDTSYFFNGASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANGASALMWLIQS 180
 Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANGASALMWLIQS 180

QY 181 TSEAAKYKFIQQIGKRVDTFLPSLAIIISLNSWSALSQIQIASTNNQFETPVVLIN 240
 Db 181 TSEAAKYKFIQQIGKRVDTFLPSLAIIISLNSWSALSQIQIASTNNQFETPVVLIN 240

QY 241 AQNRVLTITNDAGVVTNSIALLLNRNNMAAMDDDDVPMTQSGCGSYAL 289
 Db 241 AQNRVLTITNDAGVVTNSIALLLNRNNMAAMDDDDVPMTQSGCGSYAI 289

RESULT 11
 AAY87791
 ID AAY87791 standard; protein; 289 AA.
 XX AC AAY87791;
 XX 06-AUG-2003 (revised)
 DT 24-AUG-2000 (first entry)
 XX Chinese cucumber alpha-trichosanthin protein.

XX Animal RNA virus; viral coat protein; plant; male sterility; interleukin;
 KW EPO; erythropoietin; CSF; colony stimulating factor; Factor VIII; hGH;
 KW human growth hormone; melanin; insulin; vaccine;
 KW stereo specific catalysis; alpha-trichosanthin.
 XX Trichosanthin kirilowii.
 XX US6054566-A.
 XX 25-APR-2000.
 XX 07-JUN-1995; 95US-00484341.
 XX 26-FEB-1988; 88US-00160766.
 XX 26-FEB-1988; 88US-00160771.

15-JUL-1988; 88US-00219279.
 17-FEB-1989; 89US-00310881.
 05-MAY-1989; 89US-00347637.
 08-JUN-1989; 89US-00363138.
 22-OCT-1990; 90US-00600244.
 16-JAN-1991; 91US-00641617.
 26-JUN-1991; 91US-00737899.
 01-AUG-1991; 91US-00739143.
 31-JUL-1992; 92US-00923692.
 (BIOS-) BIOSOURCE TECHNOLOGIES INC.
 Garger SJ, Grill LK, Turpen TH, Grantham GL, Dawson WO;
 Turpen AM, Donson J;
 N-PSDB; AAA12393.
 Recombinant non-retroviral nucleic acid for producing proteins such as
 interleukins, melanin and vaccines, comprises subgenomic promoters linked
 to sequences coding for viral coat protein and heterologous proteins.
 Example 4; Col 61-62; 51pp; English.
 This invention describes a novel recombinant viral nucleic acid (I) from
 a non-retroviral (+) sense, single stranded animal RNA virus comprising a
 nucleic acid sequence coding for a viral coat protein regulated by a
 native subgenomic promoter and other two heterologous nucleic acid
 sequences regulated by two other subgenomic promoters. (I) is useful for
 expressing foreign genes e.g. genes inducing male sterility in plants.
 (I) is also useful for producing proteins such as interleukins, EPO
 (erythropoietin), CSF (colony stimulating factor), Factor VIII, hGH
 (human growth hormone), melatonin, insulin, vaccines etc., and enzymes that
 are useful for stereo specific catalysis of organic compounds. (I) is
 stable and transcribed systemically. The dual subgenomic promoter system
 reduces the frequency of recombination thus reducing regeneration of the
 wild type virus. This sequence represents a Chinese cucumber alpha-
 trichosanthin protein which is described in the method of the invention.
 (Updated on 06-AUG-2003 to correct OS field.)
 Sequence 289 AA;
 Query Match 98.7%; Score 1413; DB 3; Length 289;
 Best Local Similarity 98.6%; Pred. No. 1.9e-126;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 61 RSLPSGQRYALIHLYTNYADETISVAIDVTNVIYMGYRAGDTSYFFENGASATEAAKYVFK 120
 61 RSLPSGQRYALIHLYTNYADETISVAIDVTNVIYMGYRAGDTSYFFENGASATEAAKYVFK 120
 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMVLIQS 180
 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMVLIQS 180
 181 TSEAAKYKFEQIQIGKRVKDTFLPSLAISLNSWSALSQKIQIASTNNQGFETPVVLIN 240
 181 TSEAAKYKFEQIQIGKRVKDTFLPSLAISLNSWSALSQKIQIASTNNQGFETPVVLIN 240
 241 AQNRQVTITNDVAGVTSNTIALLLNRNMAAMDDVPMTQSGCGSYAL 289
 241 AQNRQVTITNDVAGVTSNTIALLLNRNMAAMDDVPMTQSGCGSYAL 289
 RESULT 12
 AAR32986
 ID AAR32986 standard; protein; 289 AA.
 XX
 AC AAR32986;
 XX

27-AUG-2003 (revised)
 25-MAR-2003 (revised)
 17-JUN-1993 (first entry)
 Encodes Chinese cucumber alpha-tricosanthin.
 Recombinant products; commercial production; fermentation; biosynthesis;
 natural products; recombinant proteins; product expression;
 protein expression; expressed proteins.
 Cucurbitaceae.
 WO9303161-A1.
 18-FEB-1993.
 31-JUL-1992; 92WO-US006359.
 01-AUG-1991; 91US-00739143.
 (DONS/) DONSON J.
 (DAWS/) DAWSON W O.
 (GRAN/) GRANTHAM G L.
 (TURP/) TURPEN T H.
 (TURP/) TURPEN A M.
 (GARG/) GARGER S J.
 (GRIL/) GRILLE L K.
 Donson J, Dawson WO, Grantham GL, Turpen TH, Turpen AM;
 Garger SJ, Grille LK;
 WPI; 1993-076518/09.
 N-PSDB; AAQ37679.
 Recombinant plant viral nucleic acids - used to express a prod., e.g.
 antibody or IL-1 in a plant.
 Example 4; Page 96; 30pp; English.
 This sequence represents Chinese cucumber alpha-tricosanthin. The coding
 sequence is inserted into a recombinant plant viral nucleic acid which is
 then used to express a recombinant product (in this case alpha-
 tricosanthin) in a plant. The plant viral sequence may be from tobacco
 mosaic, cucumber green mottle, cowpea mosaic, broad bean
 mottle, rice necrosis, geminiviruses, tomato golden mosaic, Cassava
 latent and maize streak viruses. (Updated on 25-MAR-2003 to correct PN
 field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-
 2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)
 Sequence 289 AA;
 Query Match 98.5%; Score 1410; DB 2; Length 289;
 Best Local Similarity 98.3%; Pred. No. 3.6e-126;
 Matches 284; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 61 RSLPSGQRYALIHLYTNYADETISVAIDVTNVIYMGYRAGDTSYFFENGASATEAAKYVFK 120
 61 RSLPSGQRYALIHLYTNYADETISVAIDVTNVIYMGYRAGDTSYFFENGASATEAAKYVFK 120
 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMVLIQS 180
 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSAITTLFYNNANSAASALMVLIQS 180
 181 TSEAAKYKFEQIQIGKRVKDTFLPSLAISLNSWSALSQKIQIASTNNQGFETPVVLIN 240
 181 TSEAAKYKFEQIQIGKRVKDTFLPSLAISLNSWSALSQKIQIASTNNQGFETPVVLIN 240
 241 AQNRQVTITNDVAGVTSNTIALLLNRNMAAMDDVPMTQSGCGSYAL 289
 241 AQNRQVTITNDVAGVTSNTIALLLNRNMAAMDDVPMTQSGCGSYAL 289

Db 241 AQRQVMTITNDAGVTSNTIALLLNRNMAAMDDDDVPMQTQSFQCGTYAI 289

RESULT 13

AAR07523
ID AAR07523 standard; protein; 289 AA.

XX AC AAR07523;

DT 06-FEB-1991 (first entry)

XX Alpha-Trichosanthin encoded by insert sequence from clone 12.

XX trichosanthin (TCS); Human Immunodeficiency Virus; inhibition.

XX Trichosanthin kirilowii.

XX Key Location/Qualifiers
FT Region 1. .23
FT Protein /label= signal peptide
FT /label= alpha-TCS

XX W09012097-A.

XX 18-OCT-1990.

XX 04-APR-1989; 89US-00333184.

XX 04-APR-1989; 89US-00333184.

XX (GENE-) GENELABS INC.

XX Piatek M, Chow T, Fry K;

XX WPI; 1990-334847/44.

XX N-PSDB; AAQ06351.

XX Recombinant trichosanthin protein - with selective inhibitory effect on viral expression in HIV infected T-cells or monocyte-macrophase.

XX Example; Fig 20; 102pp; English.

XX PQ30E was used as a probe to identify clones containing sequences coding for TCS in a T.kirilowii genomic library. Five clones were eventually isolated and sequenced, including clone 12. They were found to have homology to the alpha-TCS coding sequence. See also AAQ06343-Q06350

XX Sequence 289 AA;

Query Match 92.6%; Score 1326; DB 2; Length 289;
Best Local Similarity 92.4%; Pred. No. 3.9e-118;
Matches 267; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MIRELVSLILTLTLPAPVEGVSFRLSGATSSSYGVFISNRKALPNERKYDIPLL 60

Db 1 MIRELVSLILTLTLPAPVEGVSFRLSGATSSSYGVFISNRKALPYERRLYDISLL 60

Qy 61 RSSLPQSQRVALIHLYTNYADETISVAIDVTNVIYMGVAGDTSYFENGASATEAAKYVFK 120

Db 61 RSTLQGSQRVALIHLYTNYADETISVAIDVTNVIYMGVAGDTSYFENGASATEAAKYVFK 120

Qy 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

Db 121 DAQRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

Qy 181 TSEAAKYKFTFQQIGKRVKDTFLPSLAISLENSWSALSQIQIASTNNGQFSPVVLIN 240

Db 181 MSEAARYKFTFQQIGKRVKDTFLPSLAISLENSWSALSQIQIASTNNGQFSPVVLIN 240

Qy 241 AQRQVMTITNDAGVTSNTIALLLNRNMAAMDDDDVPMQTQSFQCGTYAI 289

Db 241 AQRQVMTITNDAGVTSNTIALLLNRNMAVIDDHPVMAQSFQCGSYAI 289

RESULT 14

AAR29276
ID AAR29276 standard; protein; 289 AA.

XX AC AAR29276;

DT 25-MAR-2003 (revised)

DT 13-JAN-1993 (first entry)

XX Ribosome inactivating protein deduced from pQ12.

XX TCS; alpha-trichosanthin; Radix trichosanthin; abortifacient; RIP;

XX multi-gene family; Maxim; HIV-infected human T cells;

XX human immunodeficiency virus; probe.

XX Trichosanthin kirilowii.

XX Key Location/Qualifiers
FT Peptide 1. .23
FT Protein /label= signal

XX US5128460-A.

XX 07-JUL-1992.

XX 04-APR-1990; 90US-00504775.

XX 04-APR-1989; 89US-00333184.

XX 07-SEP-1989; 89US-00404326.

XX (GENE-) GENELABS INC.

XX Piatek M, Chow TP, Fry K;

XX WPI; 1992-249485/30.

XX N-PSDB; AAQ26505.

XX Nucleic acid encoding trichosanthin protein - which can be used to inactivate ribosome(s) to inhibit protein synthesis or to inhibit HIV expression.

XX Example 7; Fig 20; 53pp; English.

XX The clone PQ30E, containing a 0.6kb insert, which was isolated from a T.kirilowii library (see AAQ26501), was used as a probe to isolate additional sequences homologous to alpha-TCS, i.e. other members of the RIP multi-gene family. Four unique cloned inserts were identified and designated pQ2, pQ3, pQ12 and pQ24. Each insert encoded a putative full length RIP protein. See AAQ26499-Q26508. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 289 AA;

Query Match 92.6%; Score 1326; DB 2; Length 289;
Best Local Similarity 92.4%; Pred. No. 3.9e-118;
Matches 267; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MIRELVSLILTLTLPAPVEGVSFRLSGATSSSYGVFISNRKALPNERKYDIPLL 60

Db 1 MIRELVSLILTLTLPAPVEGVSFRLSGATSSSYGVFISNRKALPYERRLYDISLL 60

Qy 61 RSSLPQSQRVALIHLYTNYADETISVAIDVTNVIYMGVAGDTSYFENGASATEAAKYVFK 120

Db 61 RSTLQGSQRVALIHLYTNYADETISVAIDVTNVIYMGVAGDTSYFENGASATEAAKYVFK 120

Qy 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

Db 121 DAQRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMWLIQS 180

Qy 181 TSEAAKYKFTFQQIGKRVKDTFLPSLAISLENSWSALSQIQIASTNNGQFSPVVLIN 240

QY 23 GDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLTNYADET 82
Db 1 GDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLTNYADET 60
QY 83 ISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAG 142
Db 61 ISVAIDVTNVIYMGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAG 120
QY 143 KIRENIPLGLPALDLSAITTLFYNNANSAASALMVLIOSTSEAAARYKFEIOQIGKRVDTKF 202
Db 121 KIRENIPLGLPALDLSAITTLFYNNANSAASALMVLIOSTSEAAARYKFEIOQIGKRVDTKF 180
QY 203 LPSLAIIISLENSWSALSQIOIASTNNGQFESPVLINAQNRVTITNVDAGVVTSNIAL 262
Db 181 LPSLAIIISLENSWSALSQIOIASTNNGQFESPVLINAQNRVTITNVDAGVVTSNIAL 240
QY 263 LLNRNNMAAMDDDDVPMTQSFQCGSYAL 289
Db 241 LLNRNNMAAMDDDDVPMTQSFQCGSYAI 267

Search completed: September 10, 2004, 13:58:34
Job time : 124 secs

Db 181 MSEAAARYKFEIOQIGKRVDTKFLPSLAIIISLENSWSALSQIOIASTNNGQFETPVVLIN 240
QY 241 AQQRVTITNVDAGVVTSNIALLLNRNNMAAMDDDDVPMTQSFQCGSYAL 289
Db 241 AQQRVTITNVDAGVVTSNIALLLNRNNMAVIDDHVPMAQSFQCGSYAI 289
RESULT 15
AAW21703
ID AAW21703 standard; protein; 267 AA.
XX AC AAW21703;
XX DT 25-MAR-2003 (revised)
XX DT 26-SEP-1997 (first entry)
XX DE Trichosanthin.
XX pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;
KW inactivation; eukaryotic ribosome; alpha fragment; beta fragment;
KW inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver; rat;
KW ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.
XX Trichosanthes kirilowii.
OS Trichosanthes kirilowii.

Key Location/Qualifiers
FH 139. .149
FT /note= "Position of possible insertion of internal
FT peptide linker sequence"
XX US5635384-A.
XX PD 03-JUN-1997.
XX PF 26-JAN-1995; 95US-00378761.
XX PR 11-JUN-1990; 90US-00535636.
XX PR 09-DEC-1992; 92US-00987927.
XX PA (DOWC) DOWELANCO.
XX PI Hey TD, Morgan ABR, Walsh TA;
XX DR WPI; 1997-309831/28.
XX Inactive precursor of maize ribosome-inactivating protein - also chimeric
XX ribosome-inactivating protein precursors containing internal linker
XX sequences.
XX Claim 2; Col 115-118; 121pp; English.

XX The sequences given in AAW21698-710 represent Ribosome Inactivating
XX Proteins (RIP's), which may be used in the construction of the proRIP of
XX the invention. The proRIP has a selectively removable, internal peptide
XX linker. The precursor sequence is incapable of inactivating eukaryotic
XX ribosomes, but can be converted by removal of the linker into a protein
XX having alpha and beta fragments and being capable of inactivating
XX eukaryotic ribosomes. RIPs are potent inhibitors of eukaryotic protein
XX synthesis. They possess a highly specific N-glycosidase activity which
XX cleaves the glycosidic bond of adenine 4324 of rat liver ribosomal 28S
XX RNA. RIP's selectively inhibit cellular proliferation of cells, e.g.
XX cancer cells and HIV-infected T cells. The inactive proRIP proteins make
XX it possible to provide protein synthesis inhibitors with uses in
XX practical and improved ways not before possible. The RIP can be used to
XX make cytotoxic conjugates. (Updated on 25-MAR-2003 to correct PF field.)

XX SQ Sequence 267 AA;
Query Match 92.2%; Score 1321; DB 2; Length 267;
Best Local Similarity 99.3%; Pred. No. 1e-117;
Matches 265; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 13:59:12 ; Search time 138 Seconds
(without alignments)
671.589 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MIRFLVLSLLTLFLTPA.....AAMDDVPMQTSGCGSYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1413	98.7	289	12	US-10-280-725B-4
2	1413	98.7	289	14	US-10-280-679B-4
3	1214	84.8	247	9	US-09-792-793A-39
4	1214	84.8	247	15	US-10-375-209A-39
5	1202	83.9	247	14	US-10-127-890-6
6	1064	74.3	247	9	US-09-792-793A-34
7	1064	74.3	247	15	US-10-375-209A-34
8	806.5	56.3	263	14	US-10-127-890-7
9	717.5	50.1	248	14	US-10-127-890-5
10	703	49.1	263	14	US-10-127-890-4
11	419	29.3	267	14	US-10-127-890-1
12	419	29.3	267	14	US-10-282-935-1
13	419	29.3	267	15	US-10-127-890-1
14	419	29.3	576	15	US-10-440-796-1
15	340	23.7	198	14	US-10-083-336A-3

16	340	23.7	198	14	US-10-083-336A-7	Sequence 7, Appli
17	340	23.7	199	14	US-10-083-336A-5	Sequence 5, Appli
18	340	23.7	200	14	US-10-083-336A-10	Sequence 10, Appli
19	335.5	23.4	185	14	US-10-083-336A-9	Sequence 9, Appli
20	334	23.3	188	14	US-10-083-336A-4	Sequence 4, Appli
21	334	23.3	188	14	US-10-083-336A-8	Sequence 8, Appli
22	334	23.3	189	14	US-10-083-336A-6	Sequence 6, Appli
23	334	23.3	190	14	US-10-083-336A-11	Sequence 11, Appli
24	327	22.8	251	14	US-10-282-935-3	Sequence 3, Appli
25	327	22.8	251	15	US-10-440-796-3	Sequence 3, Appli
26	327	22.8	293	9	US-09-765-527-259	Sequence 259, App
27	327	22.8	309	9	US-09-765-527-253	Sequence 253, App
28	327	22.8	332	9	US-09-765-527-251	Sequence 251, App
29	323	22.6	252	9	US-09-347-064-2	Sequence 2, Appli
30	323	22.6	252	9	US-09-347-064-8	Sequence 8, Appli
31	320	22.3	251	14	US-10-127-890-99	Sequence 99, Appli
32	320	22.3	251	14	US-10-127-890-101	Sequence 101, Appli
33	320	22.3	251	14	US-10-127-890-107	Sequence 107, Appli
34	319	22.3	251	14	US-10-127-890-110	Sequence 110, Appli
35	319	22.3	251	14	US-10-127-890-111	Sequence 111, Appli
36	319	22.3	316	14	US-10-074-596-1	Sequence 1, Appli
37	318.5	22.2	507	14	US-10-074-596-11	Sequence 11, Appli
38	318	22.2	251	9	US-09-765-527-247	Sequence 247, App
39	318	22.2	251	14	US-10-127-890-2	Sequence 2, Appli
40	318	22.2	251	14	US-10-127-890-100	Sequence 100, App
41	318	22.2	251	14	US-10-127-890-102	Sequence 102, App
42	318	22.2	251	14	US-10-127-890-103	Sequence 103, App
43	318	22.2	251	14	US-10-127-890-105	Sequence 105, App
44	318	22.2	251	14	US-10-127-890-109	Sequence 109, App
45	317	22.1	251	14	US-10-127-890-106	Sequence 106, App

ALIGNMENTS

RESULT 1

US-10-280-725B-4
; Sequence 4, Application US/10280725B
; Publication No. US20040049025A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: Recombinant Viral Nucleic Acids
; FILE REFERENCE: LSEC-0109-US02
; CURRENT APPLICATION NUMBER: US/10/280,725B
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/557,941
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 08/484,341
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/923,692
; PRIOR FILING DATE: 1992-07-31
; PRIOR APPLICATION NUMBER: 07/600,244
; PRIOR FILING DATE: 1990-10-22
; PRIOR APPLICATION NUMBER: 07/641,617
; PRIOR FILING DATE: 1991-01-16
; PRIOR APPLICATION NUMBER: 07/737,899
; PRIOR FILING DATE: 1991-07-26
; PRIOR APPLICATION NUMBER: 07/739,143
; PRIOR FILING DATE: 1991-08-01
; PRIOR APPLICATION NUMBER: 07/310,881
; PRIOR FILING DATE: 1989-02-17
; PRIOR APPLICATION NUMBER: 07/160,766
; PRIOR FILING DATE: 1988-02-26
; PRIOR APPLICATION NUMBER: 07/160,771
; PRIOR FILING DATE: 1988-02-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Chinese cucumber
US-10-280-725B-4

Query Match 98.7%; Score 1413; DB 12; Length 289;
 Best Local Similarity 98.6%; Pred. No. 3.7e-134;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTTPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 DB 1 MIRFLVLSLLILTLFTTTPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPQSQRVYALHLNLTNYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVEK 120
 DB 61 RSSLPQSQRVYALHLNLTNYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVEK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
 DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180

QY 181 TSEAAKYKFEIQOIGKRVKDTFLPSLAIISLENSWSALSQIQIASTNNGQFETPVVLIN 240
 DB 181 TSEAAKYKFEIQOIGKRVKDTFLPSLAIISLENSWSALSQIQIASTNNGQFETPVVLIN 240

QY 241 AQNRVTITNVDAAGVTSNIALLLNRNMAAMDDVPMTQSGCGSYAL 289
 DB 241 AQNRVTITNVDAAGVTSNIALLLNRNMAAMDDVPMTQSGCGSYAI 289

RESULT 2

US-10-280-679B-4
 ; Sequence 4, Application US/10280679B
 ; Publication No. US20030150019A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Large Scale Biology Corporation
 ; TITLE OF INVENTION: Monopartite RNA Virus Transformation Vectors
 ; FILE REFERENCE: LSPC-0109-US03
 ; CURRENT APPLICATION NUMBER: US/10/280,679B
 ; PENDING FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: 09/557,941
 ; PRIOR FILING DATE: 2000-04-24
 ; PRIOR APPLICATION NUMBER: 08/484,341
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: 07/923,692
 ; PRIOR FILING DATE: 1992-07-31
 ; PRIOR APPLICATION NUMBER: 07/600,244
 ; PRIOR FILING DATE: 1990-10-22
 ; PRIOR APPLICATION NUMBER: 07/641,617
 ; PRIOR FILING DATE: 1991-01-16
 ; PRIOR APPLICATION NUMBER: 07/737,899
 ; PRIOR FILING DATE: 1991-07-26
 ; PRIOR APPLICATION NUMBER: 07/739,143
 ; PRIOR FILING DATE: 1991-08-01
 ; PRIOR APPLICATION NUMBER: 07/310,881
 ; PRIOR FILING DATE: 1989-02-17
 ; PRIOR APPLICATION NUMBER: 07/160,766
 ; PRIOR FILING DATE: 1988-02-26
 ; PRIOR APPLICATION NUMBER: 07/160,771
 ; PRIOR FILING DATE: 1988-02-26
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 289
 ; TYPE: PRT
 ; ORGANISM: Chinese cucumber protein alpha-trichosanthin
 US-10-280-679B-4

Query Match 98.7%; Score 1413; DB 14; Length 289;
 Best Local Similarity 98.6%; Pred. No. 3.7e-134;
 Matches 285; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTTPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
 DB 1 MIRFLVLSLLILTLFTTTPAVEGDVSRFLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSSLPQSQRVYALHLNLTNYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVEK 120
 DB 61 RSSLPQSQRVYALHLNLTNYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVEK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180
 DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMVLQIS 180

QY 181 TSEAAKYKFEIQOIGKRVKDTFLPSLAIISLENSWSALSQIQIASTNNGQFETPVVLIN 240
 DB 181 TSEAAKYKFEIQOIGKRVKDTFLPSLAIISLENSWSALSQIQIASTNNGQFETPVVLIN 240

QY 241 AQNRVTITNVDAAGVTSNIALLLNRNMAAMDDVPMTQSGCGSYAL 289
 DB 241 AQNRVTITNVDAAGVTSNIALLLNRNMAAMDDVPMTQSGCGSYAI 289

RESULT 3

US-09-792-793A-39
 ; Sequence 39, Application US/09792793A
 ; Patent No. US20020168370A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McDonald, John R.
 ; APPLICANT: Coggin, Philip
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
 ; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
 ; FILE REFERENCE: 25020-601D
 ; CURRENT APPLICATION NUMBER: US/09/792,793A
 ; PENDING FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 93
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 39
 ; LENGTH: 247
 ; TYPE: PRT
 ; ORGANISM: Trichosanthens kirilowii
 US-09-792-793A-39

Query Match 84.8%; Score 1214; DB 9; Length 247;
 Best Local Similarity 99.6%; Pred. No. 3.9e-114;
 Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSQRVYALHLNLTNYADETI 83
 DB 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSQRVYALHLNLTNYADETI 60

QY 84 SVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVEKDMARKVTLPYSGNYERLQTAAGK 143
 DB 61 SVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVEKDMARKVTLPYSGNYERLQTAAGK 120

QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLQISSEAAKYKFEIQOIGKRVKDTFL 203
 DB 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLQISSEAAKYKFEIQOIGKRVKDTFL 180

QY 204 PSIAIISLENSWSALSQIQIASTNNGQFESPVLINAQNRVTITNVDAAGVTSNIAL 263
 DB 181 PSIAIISLENSWSALSQIQIASTNNGQFESPVLINAQNRVTITNVDAAGVTSNIAL 240

QY 264 LNRNMA 270
 DB 241 LNRNMA 247

RESULT 4

US-10-375-209A-39
 ; Sequence 39, Application US/10375209A
 ; Publication No. US20030215421A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McDonald, John R.
 ; APPLICANT: Coggin, Philip
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AN
 ; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
 ; FILE REFERENCE: 25020-601E
 ; CURRENT APPLICATION NUMBER: US/10/375,209A

; CURRENT FILING DATE: 2003-02-24
 ; NUMBER OF SEQ ID NOS: 93
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 39
 ; LENGTH: 247
 ; TYPE: PRT
 ; ORGANISM: Trichosantheus kirilowii
 US-10-375-209A-39

Query Match 84.8%; Score 1214; DB 15; Length 247;
 Best Local Similarity 99.6%; Pred. No. 3.9e-114;
 Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 83
 Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 60
 QY 84 SVAIDVTNVIYINGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
 Db 61 SVAIDVTNVIYINGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120
 QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFIQQIGKRVDTFL 203
 Db 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFIQQIGKRVDTFL 180
 QY 204 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAQNRVTITNVDAGVVTSNIAL 263
 Db 181 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAQNRVTITNVDAGVVTSNIAL 240
 QY 264 LNRNMA 270
 Db 241 LNRNMA 247

RESULT 5

US-10-127-890-6
 ; Sequence 6, Application US/10127890
 ; Publication No. US20030166196A1
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 Carroll, Stephen F.
 Studnika, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 NUMBER OF SEQUENCES: 173
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/127,890
 FILING DATE: 23-Apr-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,360
 FILING DATE: 13-MAY-1996
 APPLICATION NUMBER: PCT/US94/05348
 FILING DATE: 12-MAY-1994
 APPLICATION NUMBER: US 08/064,691
 FILING DATE: 12-MAY-1993
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 APPLICATION NUMBER: US 07/787,567

; FILING DATE: 04-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McNicholas, Janet M.
 ; REGISTRATION NUMBER: 32,918
 ; REFERENCE/DOCKET NUMBER: 200-70.P4
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/707-8889
 ; TELEFAX: 312/707-9155
 ; TELEX: 650 388-1248
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 247 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 US-10-127-890-6

Query Match 83.9%; Score 1202; DB 14; Length 247;
 Best Local Similarity 98.8%; Pred. No. 6.4e-113;
 Matches 244; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 83
 Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 60
 QY 84 SVAIDVTNVIYINGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
 Db 61 SVAIDVTNVIYINGYRAGDTSYFFNEASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120
 QY 144 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFIQQIGKRVDTFL 203
 Db 121 IRENIPGLPALDSAITTLFYNNANSAASALMVLIOSTSEAAKYKFIQQIGKRVDTFL 180
 QY 204 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAQNRVTITNVDAGVVTSNIAL 263
 Db 181 PSLAIISLNSWSALSQKIQIASTNNGQFSPVVLINAQNRVTITNVDAGVVTSNIAL 240
 QY 264 LNRNMA 270
 Db 241 LNRNMA 247

RESULT 6

US-09-792-793A-34
 ; Sequence 34, Application US/09792793A
 ; Patent No. US20020168370A1
 GENERAL INFORMATION:
 APPLICANT: McDonald, John R.
 ; APPLICANT: Coggins, Philip
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
 ; CURRENT APPLICATION NUMBER: US/09/792,793A
 ; FILE REFERENCE: 25020-601D
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 93
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 34
 ; LENGTH: 247
 ; TYPE: PRT
 ; ORGANISM: Bryonia dioica
 US-09-792-793A-34

Query Match 74.3%; Score 1064; DB 9; Length 247;
 Best Local Similarity 85.8%; Pred. No. 5.7e-99;
 Matches 212; Conservative 19; Mismatches 16; Indels 0; Gaps 0;
 QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 83
 Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPQSORYALIHLYTNYADETI 60
 QY 84 SVAIDVTNVIYINGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143

Db 61 SVAVDTNVIYNGYLAGDVSYFFNEASATEAAKFVKDAKKVTLPLPSGNYERLQTAAGK 120
QY 144 IRENIPGLPALDSAITTLFFYNNANSAASALMVLIOSTSEARYKFIEQOIGKRVDTKTEL 203
Db 121 IRENIPGLPALDSAITTLFFYNNANSAASALMVLIOSTSEARYKFIEQOIGKRVDTKTEL 180
QY 204 PSLAIIISLNSWSALSQKIQIASTNNQFESPVLINAGNQRTITNVDAVVVTSNIAL 263
Db 181 PSLATISLNSWSALSQKIQIASTNNQFESPVLINAGNQRTITNVDAVVVTSNIAL 240
QY 264 LNRNNMA 270
Db 241 LNRNNIA 247
RESULT 7
US-10-375-209A-34
; Sequence 34, Application US/10375209A
; Publication No. US20030215421A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; APPLICANT: Coggins, Philip
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
; FILE REFERENCE: 25020-60LE
; CURRENT APPLICATION NUMBER: US/10/375,209A
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Bryonia dioica
US-10-375-209A-34
Query Match 74.3%; Score 1064; DB 15; Length 247;
Best Local Similarity 85.8%; Pred. No. 5.7e-99;
Matches 212; Conservative 19; Mismatches 16; Indels 0; Gaps 0;
QY 24 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPGSORYALIHLYTNYADETI 83
Db 1 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPGSORYALIHLYTNYADETI 60
QY 84 SVAVDTNVIYNGYLAGDVSYFFNEASATEAAKFVKDAKKVTLPLPSGNYERLQTAAGK 143
Db 61 SVAVDTNVIYNGYLAGDVSYFFNEASATEAAKFVKDAKKVTLPLPSGNYERLQTAAGK 120
QY 144 IRENIPGLPALDSAITTLFFYNNANSAASALMVLIOSTSEARYKFIEQOIGKRVDTKTEL 203
Db 121 IRENIPGLPALDSAITTLFFYNNANSAASALMVLIOSTSEARYKFIEQOIGKRVDTKTEL 180
QY 204 PSLAIIISLNSWSALSQKIQIASTNNQFESPVLINAGNQRTITNVDAVVVTSNIAL 263
Db 181 PSLATISLNSWSALSQKIQIASTNNQFESPVLINAGNQRTITNVDAVVVTSNIAL 240
QY 264 LNRNNMA 270
Db 241 LNRNNIA 247
RESULT 8
US-10-127-890-7
; Sequence 7, Application US/10127890
; Publication No. US20030166196A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; Carroll, Stephen F.
; Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-9155
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-127-890-7
Query Match 56.3%; Score 806.5; DB 14; Length 263;
Best Local Similarity 63.5%; Pred. No. 6.9e-73;
Matches 165; Conservative 38; Mismatches 56; Indels 1; Gaps 1;
QY 24 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPGSORYALIHLYTNYADETI 83
Db 1 DVSPRLSGATSSSYGVFISNLRKALPNERKLYDIPLLRSSLPGSORYALIHLYTNYADETI 60
QY 84 SVAVDTNVIYNGYLAGDVSYFFNEASATEAAKFVKDAKKVTLPLPSGNYERLQTAAGK 143
Db 61 SVAVDTNVIYNGYLAGDVSYFFNEASATEAAKFVKDAKKVTLPLPSGNYERLQTAAGK 120
QY 144 IRENIPGLPALDSAITTLFFYNNANSAASALMVLIOSTSEARYKFIEQOIGKRVDTKTEL 203
Db 121 PREKIPIGLPALDSAITTLHYDSTAAGALLVLIQTAAARFKYIEQOIGERAVRDEV 180
QY 204 PSLAIIISLNSWSALSQKIQIASTNNQFESPVLINAGNQRTITNVDAVVVTSNIAL 263
Db 181 PSLATISLNSWSALSQKIQIASTNNQFESPVLINAGNQRTITNVDAVVVTSNIAL 240
QY 264 LNRNNMA 270
Db 241 LNRNNIA 247
RESULT 9
US-10-127-890-5
; Sequence 5, Application US/10127890
; Publication No. US20030166196A1

GENERAL INFORMATION:
APPLICANT: Better, Marc D.
Carroll, Stephen F.
Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 248 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-127-890-5
50.1%; Score 717.5; DB 14; Length 248;
Best Local Similarity 58.6%; Pred. No. 6.3e-64;
Matches 146; Conservative 49; Mismatches 51; Indels 3; Gaps 3;
QY 24 DVSRFLSGATSSSYGVFISNLKALPNERKLYDIPLLRSSLPQSQRVALIHLTNVADETI 83
Db 1 DVSRFLSGSSSTSYKIGDLRKLPSNGTVYNTILLSSASGASRYTLTSLSYNDGKAI 60
QY 84 SVAIDVTNVIMGYRAGDTSYFFNGASATEAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
Db 61 TVAVDVSQLVIMGYLVNSTSYFFNESDAKLASQYVFKGS-TIVTLPSYSGNYERLQTAAGK 119
QY 144 IRENIPGLPALDSAITTLTFVYNANSAASALMWLIQSTSEAKYKFIQQIGKRVDTFL 203
Db 120 IREKIPGLFALDSALTTFIHYDSTAAAPLVLQTTAEASRFKYIEGQIIBRSKNQV 179
QY 204 PSLAISLENS-WSALSQIQIASTNNQFESPVLINAQNQVTTINVDAGVTSNIAL 262
Db 180 PSLATISLENSLWSLSKIQIQAQTNNGTKPTVITDDKQQRVEINVTSKVVTKNIQ 239

QY 263 LLN-RNNMA 270
Db 240 LLNKKQNV 248
RESULT 10
US-10-127-890-4
Sequence 4, Application US/10127890
Publication No. US20030166196A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
Carroll, Stephen F.
Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-127-890-4
49.1%; Score 703; DB 14; Length 263;
Best Local Similarity 61.2%; Pred. No. 2e-62;
Matches 148; Conservative 30; Mismatches 62; Indels 2; Gaps 2;
QY 24 DVSRFLSGATSSSYGVFISNLKALPNERKLYDIPLLRSSLPQSQRVALIHLTNVADETI 83
Db 1 DVSRFLSGSSSTSYKIGDLRKLPSNGTVYNTILLSSASGASRYTLTSLSYNDGKAI 60
QY 84 SVAIDVTNVIMGYRAGDTSYFFNGASATEAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
Db 61 SVAIDVTNVIMGYRAGDTSYFFNGASATEAKYVFKDAMRKVTLPYSGNYERLQTAAGK 118

144 IRENIPGLPALDSAITTLFYNNANSAALMWLIQSTSEARFYKIEQIQKRVDTFL 203
119 IRENIDGLPALDSAITTLFYNNANSAALMWLIQSTSEARFYKIEQIQKRVDTFL 178
204 PSALISLNSWALSQKQIATSTNNQSPVVLNAQNRVTITNDAGVVTNSIAL 263
179 PNLAISLNSWALSQKQIATSTNNQSPVVLNAQNRVTITNDAGVVTNSIAL 238
264 LN 265
239 LN 240

RESULT 11
US-10-282-935-1
Sequence 1, Application US/10282935
Publication No. US20030143193A1
GENERAL INFORMATION:
APPLICANT: VITETTA, ELLEN S.
APPLICANT: GHETTIE, VICTOR F.
APPLICANT: SMALLSHAW, JOAN
APPLICANT: BALUNA, ROXANA G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF
TITLE OF INVENTION: PROTEINACEOUS COMPOUNDS
FILE REFERENCE: UTSD:884US
CURRENT APPLICATION NUMBER: US/10/282,935
CURRENT FILING DATE: 2002-10-29
PRIOR APPLICATION NUMBER: 09/538,873
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 60/126,826
PRIOR FILING DATE: 1999-03-30
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 267
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide
US-10-282-935-1

Query Match 29.3%; Score 419; DB 14; Length 267;
Best Local Similarity 36.6%; Pred. No. 1.1e-33;
Matches 93; Conservative 60; Mismatches 85; Indels 16; Gaps 7;

QY 25 VSRFLSGATSSYGVFISNLRKALPNERKL-YDIPLL--RSSLPQSQRVALIHLTNAYDE 81
DB 9 INFTTAGATVQSYTFIRAVRGLTTGADVHEIPVLPNRVGLPINQRFILVELSHAE 68
QY 82 TISVAIDVTNVYIMVGRAGDTSYFF--NGASATEAAKYVFKDMRKVTLPYSGNYERLOT 139
DB 69 SVTLALDVNTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNRVTTFAGGNYDRLEQ 128
QY 140 AAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARFYKIEQ 194
DB 129 LAGNLRNIELGNGLPEEAISALYYSTGTGTQPTLARSFIICIQMISEARFYIEGEM 188
QY 195 GKRIV--DKTFLPSLAISLNSWALSQKQIATSTNNQSPVVLNAQNRVTITND 252
DB 189 RTRIRNRRSAPDPSPVITLNSWGRSLTAIQ--ESNQGFASPIQLORNGSKFSV--YD 244
QY 253 AGVVTNSIALLLNR 266
DB 245 VSILIPITIALMVR 258

RESULT 12
US-10-127-890-1
Sequence 1, Application US/10127890
Publication No. US20030166196A1
GENERAL INFORMATION:

APPLICANT: Better, Marc D.
Carroll, Stephen F.
Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70-P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-127-890-1

Query Match 29.3%; Score 419; DB 14; Length 267;
Best Local Similarity 36.6%; Pred. No. 1.1e-33;
Matches 93; Conservative 60; Mismatches 85; Indels 16; Gaps 7;

QY 25 VSRFLSGATSSYGVFISNLRKALPNERKL-YDIPLL--RSSLPQSQRVALIHLTNAYDE 81
DB 9 INFTTAGATVQSYTFIRAVRGLTTGADVHEIPVLPNRVGLPINQRFILVELSHAE 68
QY 82 TISVAIDVTNVYIMVGRAGDTSYFF--NGASATEAAKYVFKDMRKVTLPYSGNYERLOT 139
DB 69 SVTLALDVNTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNRVTTFAGGNYDRLEQ 128
QY 140 AAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARFYKIEQ 194
DB 129 LAGNLRNIELGNGLPEEAISALYYSTGTGTQPTLARSFIICIQMISEARFYIEGEM 188
QY 195 GKRIV--DKTFLPSLAISLNSWALSQKQIATSTNNQSPVVLNAQNRVTITND 252
DB 189 RTRIRNRRSAPDPSPVITLNSWGRSLTAIQ--ESNQGFASPIQLORNGSKFSV--YD 244
QY 253 AGVVTNSIALLLNR 266

```

Db      245 VSILPIIIMVYR 258
      ::|||::|
RESULT 13
US-10-440-796-1
; Sequence 1, Application US/10440796
; Publication No. US20040009148A1
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
; TITLE OF INVENTION: SYNDROME (VLS)
; FILE REFERENCE: UTSD:603
; CURRENT APPLICATION NUMBER: US/10/440,796
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US/09/538,873
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/126,826
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-440-796-1

Query Match      29.3%; Score 419; DB 15; Length 267;
Best Local Similarity 36.6%, Pred. No. 1.1e-33;
Matches 93; Conservative 60; Mismatches 85; Indels 16; Gaps 7;

QY 25 VSFRLSGATSSYGVFISNLKALEPNERKL--YDIPLL--RSSLPGSORYALIHLYNVADE 81
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
9  INFTTAGATVQSYTNFIRAVRGLTTGADVHEIPVLPNRVGLPINOQRFILVELSNHAEL 68

QY 82 TISVAIDVTNVIYMGYRAGDTSYFF--NGASATEAAKYVFKDMRKVTLPLYSGNYERLQT 139
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69  SVTLALDVTNAYVYGVYRAGNSAYFFHPDQEDAEATHLFTDQNRVYTFAGGNYDLEQ 128

QY 140 AAGKIRENIPLGLPALDSAITLPPYNNAN-----SAASALMVLIOQTSSEAAKYKFIQQI 194
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129  LAGNLRENIELGNGPLEEALSAIYYSTGTGQTLARSFIICQMTSEAAKPYIEGEM 188

QY 195 GKRV--DKTEPLSLAIISLENSALSLSKQIASTNNGQFESPVLINAQNRVITNVVD 252
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
189  RTRIRYNRRSAPDPSVITLNSWGRSLTAIQ--ESNQAFASPIQLORNRGKFSV--YD 244

QY 253 AGVVTSNIALLLNR 266
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245  VSILPIIIMVYR 258

RESULT 14
US-10-083-336A-1
; Sequence 1, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olsson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Rigin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1

```

us-09-905-247-1.rapb

Wed Sep 15 10:32:04 2004

Search completed: September 10, 2004, 14:12:01
Job time : 139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2004, 13:50:56 ; Search time 41 Seconds
(without alignments)
678.033 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MRFVLVSLILTLFLTTFA.....AAMDVPMQTSGFGSGYAL 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1422	99.3	289	1	RLT2T
2	1377	96.2	289	2	rRNA N-glycosidase
3	1187	82.9	247	2	karasurin C - Tric
4	1182	82.5	247	2	karasurin - Mongol
5	859.5	60.0	286	1	karasurin-B - Tric
6	802.5	56.0	277	2	rRNA N-glycosidase
7	775	54.1	278	2	rRNA N-glycosidase
8	756	52.8	286	2	beta-luffin - smoo
9	753	52.6	286	2	rRNA N-glycosidase
10	673	47.0	250	2	rRNA N-glycosidase
11	663.5	46.3	245	2	luffin-b - smooth
12	419	29.3	576	1	rRNA N-glycosidase
13	394.5	27.5	564	1	ricin D precursor
14	356.5	24.9	570	2	agglutinin precurs
15	349	24.4	251	2	agglutinin I precu
16	343	24.0	562	2	abrin (clone 7.2)
17	340	23.7	528	2	abrin-c precursor
18	337.5	23.6	527	2	abrin-d precursor
19	333	23.3	528	1	abrin-a precursor
20	322	22.5	254	2	mistletoe lectin I
21	319	22.3	316	2	rRNA N-glycosidase
22	285.5	19.9	272	2	rRNA N-glycosidase
23	271.5	19.0	294	2	betavulgin - beet
24	265	18.5	313	2	rRNA N-glycosidase
25	237.5	16.6	261	2	rRNA N-glycosidase
26	215.5	15.0	310	2	antiviral protein
27	198.5	13.9	278	2	ribosome-inactivat
28	181.5	12.7	289	2	rRNA N-glycosidase
29	168	11.7	45	2	TAP-29 anti-HIV pr

ALIGNMENTS

RESULT 1

RLT2T

rRNA N-glycosidase (EC 3.2.2.22) alpha-trichosanthin precursor [validated] - Mongolian sr
N/Alternate names: alpha-TCS; type I ribosome-inactivating protein
C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)
C:Date: 30-Sep-1988 #sequence revision 26-Jan-1996 #text_change 23-Mar-2001
C:Accession: J05666; A36274; J01093; A36273; J00003
R:Shaw, P.C.; Yung, M.H.; Zhu, R.H.; Ho, W.K.K.; Ng, T.B.; Yeung, H.W.
Gene 97, 267-272, 1991
A>Title: Cloning of trichosanthin cDNA and its expression in Escherichia coli.
A:Reference number: J05666; MUID:91153657; PMID:1999291
A:Accession: J05666
A:Molecule type: mRNA
A:Residues: 1-289 <SHA>
A:Cross-references: GB:M34858; NID:g170536; PIDN:AAA34207.1; PID:g170537
A:Experimental source: tuber
R:Chow, T.P.; Feldman, R.A.; Lovett, M.; Piatak, M.
J. Biol. Chem. 265, 8670-8674, 1990
A>Title: Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I ribo
A:Reference number: A36274; MUID:90256790; PMID:2341400
A:Accession: A36274
A:Molecule type: DNA
A:Residues: 1-233, 'T', 235-245, 'M', 248-289 <CHO>
A:Cross-references: GB:J05434; NID:g170534; PIDN:AAA34206.1; PID:g170535
R:Zheng, H.G.; Wang, B.; Shao, P.Z.; Yang, X.R.
Acta Genet. Sin. 21, 42-51, 1994
A>Title: Cloning and DNA sequencing of the gene encoding Trichosanthin.
A:Reference number: J01093; MUID:94271613; PMID:8003348
A:Accession: J01093
A:Molecule type: DNA
A:Residues: 1-72, 'V', 74-90, 'S', 92-233, 'T', 235-267, 'D', 269-289 <ZHE>
A:Cross-references: GB:S70176; NID:G547148; PIDN:AAH31048.1; PID:G547149
R:Collins, E.J.; Robertus, J.D.; LoPresti, M.; Stone, K.L.; Williams, K.R.; Wu, P.; Hwang
J. Biol. Chem. 265, 8665-8669, 1990
A>Title: Primary amino acid sequence of alpha-trichosanthin and molecular models for abri
A:Reference number: A36273; MUID:90256789; PMID:2341399
A:Accession: A36273
A:Molecule type: protein
A:Residues: 24-270 <COL>
R:Wang, Y.; Qian, R.Q.; Gu, Z.W.; Jin, S.W.; Zhang, L.Q.; Xia, Z.X.; Tian, G.Y.; Ni, C.Z.
Pure Appl. Chem. 58, 789-798, 1986
A>Title: Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application.
A:Reference number: J00003
A:Accession: J00003
A:Molecule type: protein
A:Residues: 24-56, 'L', 58-59, 'I', 61-71, 'I', 73-81, 85-86, 'L', 88-92, 'DAGLPRNAVL', 93-142, 'GL',
R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.
submitted to the Brookhaven Protein Data Bank, July 1994
A:Reference number: A67091; PDB:1MRJ
A:Contents: annotation: X-ray crystallography, 1.6 angstroms, with adenine, residues 24-;
R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.

submitted to the Brookhaven Protein Data Bank, July 1994
A;Reference number: A67092; PDB:IMRK
A;Contents: annotation; X-ray crystallography, 1.6 angstroms, with formycin, residues 24-27
R;Xiong, J.P.; Xia, Z.X.; Wang, Y.
submitted to the Brookhaven Protein Data Bank, December 1994
A;Reference number: A66711; PDB:1TCS
A;Contents: annotation; X-ray crystallography, 1.7 angstroms, with NADPH, residues 24-27
R;Xiong, J.P.; Xia, Z.X.; Wang, Y.
Nat. Struct. Biol. 1, 695-700, 1994
A;Title: Crystal structure of trichosanthin-NADPH complex at 1.7 Angstroms resolution re
A;Reference number: A58622; MUID:95360714; PMID:7634073
A;Contents: annotation; X-ray crystallography, 1.7 angstroms
C;Comment: Alpha-trichosanthin has been used to induce abortions.
C;Genetics:
A;Gene: tcs
C;Function:
A;Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA then
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C;Keywords: abortifacient; glycosidase; hydrolase; root; toxin
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-270/Product: trichosanthin alpha #status experimental <WAT>
F;27-266/Domain: rRNA N-glycosidase homology <RNG>
F;271-289/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;93,183,186/Active site: Tyr, Glu, Arg #status predicted

Query Match 99.3%; Score 1422; DB 1; Length 289;
Best Local Similarity 99.3%; Pred. No. 6.2e-108;
Matches 287; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDSVRLSGATSSSYGVFISNLRKALPNERKLYDIPL 60
Db 1 MIRFLVLSLLILTLFTTPAVEGDSVRLSGATSSSYGVFISNLRKALPNERKLYDIPL 60

QY 61 RSSLPGSQRYALIHLYTNVADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120
Db 61 RSSLPGSQRYALIHLYTNVADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMVLIO 180
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMVLIO 180

QY 181 TSEAAKYKFIQOIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNGQFESPVLIN 240
Db 181 TSEAAKYKFIQOIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNGQFESPVLIN 240

QY 241 AQNRVTITNDVAGVVTSNIALLLNRNNMAAMDDVPMTQSGCGSYAL 289
Db 241 AQNRVTITNDVAGVVTSNIALLLNRNNMAAMDDVPMTQSGCGSYAI 289

RESULT 2
JC5606
karasurin C - Trichosanthes kirilowii var. japonica
N;Contains: karasurin A
C;Species: Trichosanthes kirilowii var. japonica
C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 19-Jul-2002
C;Accession: JC5606; JC5033
R;Mizukami, H.; Iida, K.; Kondo, T.; Ogiwara, Y.
Biol. Pharm. Bull. 20, 711-713, 1997
A;Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating prote
A;Reference number: JC5606; MUID:97356562; PMID:9212998
A;Accession: JC5606
A;Molecule type: DNA
A;Residues: 1-289 <MIZ>
A;Cross-references: DDBJ:AB006666; NID:92329830; PIDN:BA21786.1; PID:92329831
R;Kondo, T.; Mizukami, H.; Takeda, T.; Ogiwara, Y.
Biol. Pharm. Bull. 19, 1485-1489, 1996
A;Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka
A;Reference number: JC5032; MUID:97108848; PMID:8951169
A;Accession: JC5033
A;Status: preliminary
A;Molecule type: protein
A;Residues: 22-270 <KON>

C;Comment: This protein is a ribosome-inactivating protein and exhibit cytotoxic, abortif
C;Comment: This protein belongs to type I ribosomal-inactivating proteins which catalytic
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
F;22-270/Product: karasurin C #status predicted <MAC>
F;24-270/Product: karasurin A #status predicted <MAC>
F;27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match 96.2%; Score 1377; DB 2; Length 289;
Best Local Similarity 95.8%; Pred. No. 2.8e-104;
Matches 277; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDSVRLSGATSSSYGVFISNLRKALPNERKLYDIPL 60
Db 1 MIRFLVLSLLILTLFTTPAVEGDSVRLSGATSSSYGVFISNLRKALPNERKLYDIPL 60

QY 61 RSSLPGSQRYALIHLYTNVADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120
Db 61 RSSLPGSQRYALIHLYTNVADETISVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMVLIO 180
Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMVLIO 180

QY 181 TSEAAKYKFIQOIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNGQFESPVLIN 240
Db 181 TSEAAKYKFIQOIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNGQFESPVLIN 240

QY 241 AQNRVTITNDVAGVVTSNIALLLNRNNMAAMDDVPMTQSGCGSYAL 289
Db 241 AQNRVTITNDVAGVVTSNIALLLNRNNMAAMDDVPMTQSGCGSYAI 289

RESULT 3
JU0393
karasurin - Mongolian snake-gourd
C;Species: Trichosanthes kirilowii (Mongolian snake-gourd)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 12-Apr-1995
C;Accession: JU0393; PS0163
R;Toyokawa, S.; Takeda, T.; Kato, Y.; Wakabayashi, K.; Ogiwara, Y.
Chem. Pharm. Bull. 39, 1244-1249, 1991
A;Title: The complete amino acid sequence of an abortifacient protein, karasurin.
A;Reference number: JU0393; MUID:92005921; PMID:1914000
A;Accession: JU0393
A;Molecule type: protein
A;Residues: 1-247 <TOY>
A;Note: a sequence which lacks Ala-247 is also shown in this publication
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C;Keywords: abortifacient
F;4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 82.9%; Score 1187; DB 2; Length 247;
Best Local Similarity 97.2%; Pred. No. 5.8e-89;
Matches 240; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLRSLSQRYALIHLYTNVADETI 83
Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLRSLSQRYALIHLYTNVADETI 60

QY 84 SVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
Db 61 SVAIDVTNVIYMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120

QY 144 IRENIPGLPALDSAITTLFYNNANSAALMVLIOSTSEAAKYKFIQOIGKRVDTFL 203
Db 121 IRENIPGLPALDSAITTLFYNNANSAALMVLIOSTSEAAKYKFIQOIGKRVDTFL 180

QY 204 PSIAISLNSWSALSQKIQIASTNNGQFESPVLINAQNRVTITNDVAGVVTSNIAL 263
Db 181 PSIAISLNSWSALSQKIQIASTNNGQFESPVLINAQNRVTITNDVAGVVTSNIAL 240

QY 264 LNENNMA 270
Db 241 LNENNMA 247

RESULT 4

JCS032
 karasurin-B - Trichosanthes kirilowii var. japonica
 C;Species: Trichosanthes kirilowii var. japonica
 C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 23-May-1997
 A;Accession: JCS032
 R;Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.
 Biol. Pharm. Bull. 19, 1485-1489, 1996
 A;Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and karasurin-A
 A;Reference number: JCS032; MUID:97108848; PMID:8951169
 A;Accession: JCS032
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-247 <KON>
 C;Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyze the inactivation of 28S rRNA N-glycosidase
 F;243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 82.5%; Score 1182; DB 2; Length 247;
 Best Local Similarity 96.8%; Pred. No. 1.5e-88;
 Matches 239; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 24 DVSFRLSGATSSSYGVFISNLRKALPNERKLDIPLRLSLPGSQRYALHLTNVADETI 83
 Db 1 DVSFRLSGATSSSYGVFISNLRKALPNERKLDIPLRLSLPGSQRYALHLTNVADETI 60
 QY 84 SVAIDVTNYYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
 Db 61 SVAIDVTNYYIMGYRAGDTSYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 120
 QY 144 IRENTPLGLPALDSAITTLFYNNANSASALMVLIOSTSEAAKYFIEQIQKRVDTKPL 203
 Db 121 IRENTPLGLPALDSAITTLFYNNANSASALMVLIOSTSEAAKYFIEQIQKRVDTKPL 180
 QY 204 PSALISLNSWSALSCKOIQIASTNNGOPESPVVLINAGNQRTVITNDVAGVVTNNIAL 263
 Db 181 PSALISLNSWSALSCKOIQIASTNNGOPETFPVLINAGNQRTVITNDVAGVVTNNIAL 240
 QY 264 LNRNMA 270
 Db 241 LNRNMA 247

RESULT 5

RUPUGG
 rRNA N-glycosidase (EC 3.2.2.22) alpha-momorcharin precursor [validated] - balsam pear
 N;Alternate names: agglutinin; momordin-A; ribosome-inactivating protein momorcharin alp
 C;Species: Momordica charantia (balsam pear, bitter gourd)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 15-Sep-2000
 A;Accession: S14273; A61318; S16490; JN0628; S01670
 R;Ho, W.K.K.; Liu, S.C.; Shaw, P.C.; Yeung, H.W.; Ng, T.B.; Chan, W.Y.
 Biochim. Biophys. Acta 1088, 311-314, 1991
 A;Title: Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating protein.
 A;Reference number: S14273; MUID:91159486; PMID:2001404
 A;Accession: S14273
 A;Molecule type: mRNA
 A;Residues: 1-286 <HOW>
 A;Cross-references: EMBL:X57682; NID:gl9527; PID:CAA40869.1; PID:gl9528
 R;Li, S.S.L.,

Experientia 36, 524-527, 1980
 A;Title: Purification and partial characterization of two lectins from Momordica charantia
 A;Reference number: A61318; MUID:80201763; PMID:7379938
 A;Accession: A61318
 A;Molecule type: protein
 A;Residues: 24-50 <LTA>
 A;Note: as a lectin shows agglutinating activity for type-O red blood cells
 R;Montecucchi, P.C.; Lazzerini, A.M.; Barbieri, L.; Stalpe, F.; Sorla, M.; Lappi, D.
 Int. J. Pept. Protein Res. 33, 263-267, 1989
 A;Title: N-terminal sequence of some ribosome-inactivating proteins.
 A;Reference number: S16331; MUID:89326691; PMID:2753596
 A;Accession: S16490

A;Molecule type: protein
 A;Residues: 24-68 'X', 70 <MON>
 R;Minami, Y.; Funatsu, G.
 Biosci. Biotechnol. Biochem. 57, 1141-1144, 1993
 A;Title: The complete amino acid sequence of momordin-a, a ribosome-inactivating protein
 A;Reference number: JN0628; MUID:93372485; PMID:7763984
 A;Accession: JN0628
 A;Molecule type: protein
 A;Residues: 24-107, 'Q', 109-123, 125-147, 'L', 149-154, 'I', 156-205, 'I', 207-208, 'L', 210-214, 'I'
 A;Experimental source: seed
 R;Ren, J.; Wang, Y.; Dong, Y.; Stuart, D.I.
 submitted to the Brookhaven Protein Data Bank, January 1994
 A;Reference number: A52272; PDB:1AHC
 A;Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-269
 R;Husain, J.; Tickle, I.J.; Wood, S.P.
 submitted to the Brookhaven Protein Data Bank, March 1994
 A;Reference number: A52385; PDB:1MOM
 A;Contents: annotation; X-ray crystallography, 2.16 angstroms, residues 24-86, 'L', 88-269
 R;Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.
 submitted to the Brookhaven Protein Data Bank, July 1994
 A;Reference number: A67089; PDB:1MRH
 A;Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-77, 'R', 79-132, 'I'
 C;Function:

A;Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA there
 C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
 C;Keywords: glycoprotein; glycosidase; hydrolase; lectin; seed; toxin
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-23/Domain: amino-terminal propeptide #status predicted <PRO>
 F;24-269/Product: rRNA N-glycosidase alpha-momorcharin #status experimental <MAT>
 F;27-266/Domain: rRNA N-glycosidase homology <RNG>
 F;270-286/Domain: carboxyl-terminal propeptide #status predicted <CTP>
 F;93.183,186/Active site: Tyr, Glu, Arg #status predicted
 F;250/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 60.0%; Score 859.5; DB 1; Length 286;
 Best Local Similarity 62.9%; Pred. No. 2.7e-62;
 Matches 178; Conservative 40; Mismatches 64; Indels 1; Gaps 1;

QY 1 MIREVLVSLILITFLITTPAVEGVDSFRLSGATSSSYGVFISNLRKALPNERKLDIPL 60
 Db 1 MSRFVLSFLILAIFLPGSIYKGVDSFRLSGADPRSYGMFKDLRNALPFREKVNIPLL 60
 QY 61 RSSLPGSORYALIHILTNVADETISVAIDVTNYYIMGYRAGDTSYFFNGASATEAAKYVFK 120
 Db 61 LPSVSGAGRYLMLHFLFYDGRKTIIVADVNTNYYIMGYRAGDTSYFFNEPAELASQYVFR 120
 QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSASALMVLIOQ 180
 Db 121 DARRKVTLPYSGNYERLQTAAGKPREKIPGLPALDSAITTLHYDSTAAAGALLVLIQT 180
 QY 181 TSEARVYKFTQQIGKRVDTKPLSLAISLNSWSALSCKOIQIASTNNGOPESPVVLIN 240
 Db 181 TAEARVYKFTQQIGKRVDTKPLSLAISLNSWSALSCKOIQIASTNNGOPESPVVLIN 240
 QY 241 AQONRVITTNVADGVTSNIALLRNNMAAMD-DYPMQTQSF 282
 Db 241 NKGNRVOITNVTSKVTSNIQLLNTNRNIAEGONGDVSTTHGF 283

RESULT 6

S22494
 rRNA N-glycosidase (EC 3.2.2.22) alpha-luffin precursor - smooth loofah
 N;Alternate names: protein synthesis inhibitor; ribosome-inactivating protein luffin-A
 C;Species: Luffa cylindrica (smooth loofah)
 C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 20-Aug-1999
 A;Accession: S22494; S26390; JH0202; A32542
 R;Kataoka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Koiwai, A.
 Plant Mol. Biol. 18, 1199-1202, 1992
 A;Title: Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-inactivating prote
 A;Reference number: S22494; MUID:92288316; PMID:1600156
 A;Accession: S22494
 A;Molecule type: mRNA
 A;Residues: 1-277 <KAT>

A;Cross-references: EMBL:X62371; NID:g19145; PIDN:CAA44229.1; PID:g19146
R;Islam, M.R.; Nishida, H.; Funatsu, G.
Agric. Biol. Chem. 54, 2967-2978, 1990
A;Title: Complete amino acid sequence of luffin-a, a ribosome-inactivating protein from
A;Reference number: S26390; MUID:91248471; PMID:1368651
A;Accession: S26390
A;Molecule type: protein
A;Residues: 21-53, 'L', 55, 'I', 57-86, 'SQL', 90-154, 'L', 156-157, 'I', 159-173, 'L', 175-209, 'SL'
R;Islam, M.R.; Nishida, H.; Funatsu, G.
Agric. Biol. Chem. 54, 1343-1345, 1990
A;Title: Complete amino acid sequence of luffin-a, a ribosome-inactivating protein from
A;Reference number: JH0202; MUID:91197482; PMID:1368623
A;Accession: JH0202
A;Molecule type: protein
A;Residues: 21-53, 'L', 55, 'I', 57-86, 'SQL', 90-154, 'L', 156-157, 'I', 159-173, 'L', 175-209, 'SL'
A;Experimental source: seed
R;Ramakrishnan, S.; Enghid, J.J.; Bryant Jr., H.L.; Xu, F.J.
Biochem. Biophys. Res. Commun. 160, 509-516, 1989
A;Title: Characterization of a translation inhibitory protein from Luffa aegyptiaca.
A;Reference number: A32542; MUID:89246493; PMID:2719679
A;Accession: A32542
A;Status: preliminary
A;Molecule type: protein
A;Residues: 21-28, 'G', 30, 'X', 32-34, 'K', 36-40 <RAM>
C;Function:
A;Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA then
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C;Keywords: glycoprotein; glycosidase; hydrolase; seed; toxin
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-267/Domain: rRNA N-glycosidase alpha-luffin #status experimental <MAT>
F;224-262/Domain: rRNA N-glycosidase homology <RNG>
F;268-277/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;48,53,97,104,225,246/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;90,179,182/Active site: Tyr, Glu, Arg #status predicted

Query Match 56.0%; Score 802.5; DB 2; Length 277;
Best Local Similarity 60.3%; Pred. No. 1.1e-57;
Matches 167; Conservative 45; Mismatches 60; Indels 5; Gaps 3;

QY 1 MIRFLVLSLLILTLFTTTPAVEGVDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MKRFTV--LILAIFVAASTVEADVRFSLSGSSSTSYSGKFIGDLRKALPNSGTVYNTLL 57

QY 61 RSSLPGSQRYALHILTNVADETISVAIDVTNVIYNGYAGDTSYFENGASATEAAKYVFX 120
DB 58 ISSASGASRYTLMTLSNYDGAITVAVDVTNVIYNGYLVNSTSYFFNDSADKLASQYVFX 117

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASALMVLIOQ 180
DB 118 GS-TIVTLPSYNGYERLQTAAGKIRENIPGLPALDSAITTLFFHYDSTAATAAFVLIQT 176

QY 181 TSEAAKYKIEQIQIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNQGFESPVLIN 240
DB 177 TAEASRFKIEQIIERISKNQVPSLATISLNSWSALSQKIQIQAOTNNGTFTKPVVITD 236

QY 241 AQNQRTVITNVDAGVVTNSIALLLN-RNNMAAMDDV 276
DB 237 DKGQRVEITNVTSKVVTNKIQLLNKQNVAAFDV 273

RESULT 7
S23519
beta-luffin - smooth loofah
C;Species: Luffa cylindrica (smooth loofah)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Aug-1999
C;Accession: S23519; MUID:92353400; PMID:1643290
R;Kataoka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Koiwai, A.
Plant Mol. Biol. 19, 887-889, 1992
A;Title: Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-inactivating
A;Reference number: S23519; MUID:92353400; PMID:1643290
A;Accession: S23519
A;Molecule type: mRNA
A;Residues: 1-278 <KAT>

A;Cross-references: EMBL:X62372; NID:g19149; PIDN:CAA44230.1; PID:g19150
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
F;26-264/Domain: rRNA N-glycosidase homology <RNG>

Query Match 54.1%; Score 775; DB 2; Length 278;
Best Local Similarity 58.6%; Pred. No. 1.9e-55;
Matches 163; Conservative 47; Mismatches 64; Indels 4; Gaps 3;

QY 1 MIRFLVLSLLILTLFTTTPAVEGVDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 59
DB 1 MNRFTFLSLLILILIAFT--VEGANVSFLSGADSKSYSKFITALRKALPSKESVNIPL 58

QY 60 LRSLPGSQRYALHILTNVADETISVAIDVTNVIYNGYAGDTSYFENGASATEAAKYVFX 119
DB 59 LPLSASGASRYTLMTLSNYDGAITVAIDVTNVIYNGYLVNSTSYFFNDSADKLASQYVFX 118

QY 120 KDMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASALMVLIOQ 179
DB 119 KGS-TIVTLPSYNGYERLQTAAGKIRENIPGLPALDSAITTLFFHYDSTAATAAFVLIQT 177

QY 180 STSEAAKYKIEQIQIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNQGFESPVLIN 239
DB 178 TTAASRFKIEQIIERIPKNEVPSPALSLNSWSALSQKIQIQAOTNNGTFTKPVVIT 237

QY 240 NAQNRVTITNVDAGVVTNSIALLLN-RNNMAAMDDV 277
DB 238 DKGQRVEITNVTSKVVTNKIQLLNKQNVAAFDV 275

RESULT 8
S25560
rRNA N-glycosidase (EC 3.2.2.22) momordin II - balsam apple
C;Species: Momordica balsamina (balsam apple)
C;Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 20-Aug-1999
C;Accession: S25560
R;Ortigao, M.; Better, M.
Nucleic Acids Res. 20, 4662, 1992
A;Title: Momordin II, a ribosome inactivating protein from Momordica balsamina, is homologous to
A;Reference number: S25560; MUID:93027170; PMID:1408771
A;Accession: S25560
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-286 <ORT>
A;Cross-references: EMBL:Z12175; NID:g19525; PIDN:CAA78166.1; PID:g19526
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C;Keywords: glycosidase; hydrolase
F;27-264/Domain: rRNA N-glycosidase homology <RNG>

Query Match 52.8%; Score 756; DB 2; Length 286;
Best Local Similarity 59.2%; Pred. No. 6.8e-54;
Matches 157; Conservative 37; Mismatches 69; Indels 2; Gaps 2;

QY 1 MIRFLVLSLLILTLFTTTPAVEGVDSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MYKLLLSFLIIITAIIGVPTAKGVNFDLSTATYTKFIEDFRATLFPFHKVYDIPLL 60

QY 61 RSSLPGSQRYALHILTNVADETISVAIDVTNVIYNGYAGDTSYFENGASATEAAKYVFX 120
DB 61 YSTISDSRRFILLDTSVAYETISVAIDVTNVIYNGYLVNSTSYFFNDSADKLASQYVFX 119

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASALMVLIOQ 180
DB 120 -GTRKITLPTNGYENLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASALMVLIOQ 178

QY 181 TSEAAKYKIEQIQIGKRVDTFLPSLAISLNSWSALSQKIQIASTNNQGFESPVLIN 240
DB 179 TAEAAKFYKIEHVAKYVATNFKENLAISLNSWSALSQKIQIQAOTNNGTFTKPVVIT 238

QY 241 AQNQRTVITNVDAGVVTNSIALLLN 265
DB 239 PTGERFQVITNVDSDVVKGNKILLN 263

A;Residues: 303-325,'F',327-330,'T',332-361,'D',363-373,'G',375-403,'T',405-551,'V',553-
 C;Comment: This protein has strong agglutinating activity and weak cytotoxicity compared
 C;Superfamily: ricin; rRNA N-glycosidase homology
 C;Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-290/Product: agglutinin chain A #status predicted <ACH>
 F;35-281/Domain: rRNA N-glycosidase homology <RNG>
 F;303-564/Product: agglutinin chain B #status experimental <BCH>
 F;319-361,362-402,405-443,450-485,489-528,531-564/Region: 40-residue repeats
 F;34,259/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;104,147,231,232/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F;200,203/Active site: Glu, Arg #status predicted
 F;282-306,322-341,365-382,453-466,492-509/Disulfide bonds: #status predicted
 F;324,337,348/Binding site: N-acetylglucosamine (Asp, Gln, Asn) #status predicted
 F;397,437/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F;536,557/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 27.5%; Score 394.5; DB 1; Length 564;
 Best Local Similarity 36.6%; Pred. No. 3.8e-24;
 Matches 93; Conservative 55; Mismatches 89; Indels 17; Gaps 8;

QY 25 VSRSLGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSSLPGRVORVALIHLTNVADE 81
 Db 33 INFTTADATVESYTFIRAVRSHLTTCGADVRLPVRVGLPIQRFILVLSNHAEL 92
 QY 82 TTSVALDVNTVIMYGRAGDTSYFF--NGASATEAAYVFKDAMRKVTLPLYSNGYERLQT 139
 Db 93 SVTLALDVNTAVYVGRAGNSAVFFPHDQEDAEATHLFTDVQNSFTFAFGNGYDRLEQ 152
 QY 140 AAGKIRENIPGLPALDSAITTLFYNA-----NSASALMVLIOQTSAAAYKFTQOI 194
 Db 153 LGG-LRENELGTGPLEDAISALYVSTCTGPTILARSFMVCIQMISEARFQYIEGEM 211
 QY 195 GRRV--DKTFLPSLAISLNSWALSLSKQIQTASTNNGQFSPVVLINAOQRVITITVD 252
 Db 212 RTRIRNRSAPDPVSITLNSWGRSLTAIQ--ESNQGAFAPSIQLRNGSKFNV--YD 267
 QY 253 AGVTSNTALLNR 266
 Db 268 VSLIPILALMYR 281

RESULT 14
 S62627
 agglutinin I precursor - European elder
 C;Species: Sambucus nigra (European elder)
 C;Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 20-Aug-1999
 C;Accession: S62627; S62619
 R;van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.
 Eur. J. Biochem. 235, 128-137, 1996
 A;Title: The NeuAc(alpha-2,6)-Gal/GalNAc-binding lectin from elderberry (Sambucus nigra)
 A;Reference number: S62619; MUID:96202926; PMID:8631319
 A;Accession: S62627
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-570 <V>
 A;Cross-references: EMBL:U07122; NID:g1141772; PIDN:AAC49158.1; PID:g1141773
 A;Accession: S62619
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 29-39;309-319 <VA2>
 C;Superfamily: ricin; rRNA N-glycosidase homology
 F;137-283/Domain: rRNA N-glycosidase homology <RNG>

Query Match 24.9%; Score 356.5; DB 2; Length 570;
 Best Local Similarity 32.5%; Pred. No. 4.7e-21;
 Matches 103; Conservative 58; Mismatches 115; Indels 41; Gaps 11;

QY 5 LVLSLLILTLF-----TTPAYGDSVSRSLGATSSSYGVFISNLR-KALP 49
 Db 3 LVAKLLVLAVALICGLIHGALHPRVTPPYVPSVFNLTGA--DTYEPFLRAQKQVIL 60
 QY 50 NERKLYDIPLI--RSSLPGRVORVALIHLTNVADETISVAIDVTNVIIMYGRAGDTSYFFN 107

Db 61 GNHTAPDLVLPESQSDSNRFVLPVPLTNPSGDTVLADVNLYVFASSNKGKSYFFS 120
 QY 108 GASATEAAKVFVKDAMRKVTLPLYSNGYERLQTAAGKIRENIPGLPALDSAITTLFYNA 167
 Db 121 GSTAVORDN-LFVDITQEE-ELNFTGNVTSLEKQVGRVYIPLGPKSLDQAISLRVYTL 178
 QY 168 NSA-----ASALMVLIOQTSAAAYKFTQOIQKRV--KTFPLPSLAISLNSWSALS 220
 Db 179 TAGDTPLARGLLIVTQMSEAAAFRIELRIRTSITDASEFTPDLLMLSMENNWSMSS 238
 QY 221 QIQIATSTNNGQFSPVVLINAOQRVITITVDAGVTSNTALL--NRRNMA 270
 Db 239 EIQAQAP-GGIFAGVQLRDERNNSIEVTFRLFELTYTAVLLYGCAPVTSSSYNN-- 295
 QY 271 AMDDVDVMTQSPGCGSY 287
 Db 296 AIDAIKMPVFRGGEY 312

RESULT 15
 C39761
 abrin (clone 7.2) precursor - Indian licorice (fragment)
 N;Contains: rRNA N-glycosidase (EC 3.2.2.22)
 C;Species: Abrus precatorius (Indian licorice)
 C;Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
 C;Accession: C39761; S14471
 R;Evensen, G.; Mathiesen, A.; Sundan, A.
 J. Biol. Chem. 266, 6848-6852, 1991
 A;Title: Direct molecular cloning and expression of two distinct abrin A-chains.
 A;Reference number: A39761; MUID:91201329; PMID:2016300
 A;Accession: C39761
 A;Molecule type: DNA
 A;Residues: 1-251 <EVE>
 R;Evensen, G.; Mathiesen, A.; Sundan, A.
 submitted to the EMBL Data Library, October 1990
 A;Description: Direct molecular cloning of two distinct abrin A-chains.
 A;Reference number: S14471
 A;Accession: S14471
 A;Molecule type: DNA
 A;Residues: 'M',1-251 <EV2>
 A;Cross-references: EMBL:X54872; NID:g16088; PIDN:CAA38654.1; PID:g16089
 C;Superfamily: ricin; rRNA N-glycosidase homology
 C;Keywords: duplication; glycosidase; hydrolase; lectin; toxin
 F;1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>
 F;7-246/Domain: rRNA N-glycosidase homology <RNG>
 F;74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F;164,167/Active site: Glu, Arg #status predicted

Query Match 24.4%; Score 349; DB 2; Length 251;
 Best Local Similarity 38.6%; Pred. No. 6.1e-21;
 Matches 95; Conservative 41; Mismatches 98; Indels 12; Gaps 8;

QY 25 VSRSLGATSSSYGVFISNLRKALPNERKLYDIPLL--SSLPGSRVALIHLTNVADET 82
 Db 5 IKFSTEGATSSYKQFTEALBERL--RGGLIHDPVLRDPTVEERNRITVELSNERES 63
 QY 83 ISVAIDVTNVIIMYGRAGDTSYFFNGASATEAAYVFKDAMRKVTLPLYSNGYERLQTAAG 142
 Db 64 IEVGIDVTNVAWYVAGSQSYFLRDAPAS--ASTVLTGTQR--YSLRFDGSGYDLERWAH 121
 QY 143 KIRENIPGLPALDSAITTL--FYNNASASALMVLIOQTSAAAYKFTQOIQK--RV 198
 Db 122 QTRQISLGLQALTHAISFLFSGASNDDEKARTLITVIQMASEAAARYISNRVGSIRT 181
 QY 199 DKTEPLSLAISLNSWALSLSKQIQTASTNNGQFSPVVLINAOQRVITITVDAGVTS 258
 Db 182 GTAQPPDAMLSLENNWDLNSRGVQ--ESVQDTPFNAVTLERRVNNQPVIVDSLTHQSV-A 238
 QY 259 NIALLL 264
 Db 239 VLALML 244

Search completed: September 10, 2004, 14:01:55
Job time : 42 secs

Db 61 RSTLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180

Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180

QY 181 TSEAAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQGFSPVVLIN 240

Db 181 TSEAAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQGFSPVVLIN 240

QY 241 AQNRQVTITNDAGVVTNSIALLLNNRNNMAAMDDDDVPMTQSFCCGSYAI 289

Db 241 AQNRQVTITNDAGVVTNSIALLLNNRNNMAAMDDDDVPMTQSFCCGSYAI 289

RESULT 2

Q94KE4 PRELIMINARY; PRT; 289 AA.

AC Q94KE4 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 25, Last annotation update)

DE Trichosanthin precursor (EC 3.2.2.22) (rRNA N-glycosidase).

GN TCS.

OS Trichosanthes kirilowii (Mongolian snake-gourd).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.

OX NCBI_TaxID=3677;

RN [1]

SEQUENCE FROM N.A.

RA Yuan H., Wang L., Wang Y., An C., Chen Z.;

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE

CC SPECIFIC ADENOSINE ON THE 28S RRNA.

CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

DR EMBL; AF367252; AAK52960.1; --

DR PIR; J05032; J05032.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.

DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.

DR GO; GO:0015070; F:toxin activity; IEA.

DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.

DR InterPro; IPR001574; RIP.

DR Pfam; PF00161; RIP; 1.

DR PRINTS; PR00396; SHIGARICIN.

DR PROSITE; PS00275; SHIGA_RICIN; 1.

KW Hydrolase; Signal; Toxin.

FT SIGNAL 1 23

FT CHAIN 24 270

FT TRICHOSANTHIN

SQ SEQUENCE 289 AA; 31706 MW; A6D5602549CA5657 CRC64;

Query Match 99.0%; Score 1418; DB 10; Length 289;

Best Local Similarity 98.6%; Pred. No. 4.2e-104;

Matches 285; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

Db 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFK 120

Db 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180

Db 121 DSMRKITLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180

QY 181 TSEAAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQGFSPVVLIN 240

Db 181 TSEAAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQGFSPVVLIN 240

QY 241 AQNRQVTITNDAGVVTNSIALLLNNRNNMAAMDDDDVPMTQSFCCGSYAI 289

Db 241 AQNRQVTITNDAGVVTNSIALLLNNRNNMAAMDDDDVPMTQSFCCGSYAI 289

RESULT 3

Q41216 PRELIMINARY; PRT; 289 AA.

AC Q41216 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-NOV-1996 (TrEMBLrel. 25, Last annotation update)

DE Trichosanthin (EC 3.2.2.22) (rRNA N-glycosidase).

GN TRICHOSANTHIN, TCS.

OS Trichosanthes kirilowii (Mongolian snake-gourd).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.

OX NCBI_TaxID=3677;

RN [1]

SEQUENCE FROM N.A.

RA MEDLINE=94271613; PubMed=8003348;

RA Zheng H., Wang B., Shaw P., Yeung H.;

RT "[Cloning and DNA sequencing of the gene encoding trichosanthin].";

RL I Chuan Hsueh Pao 21:42-51(1994).

CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE

CC SPECIFIC ADENOSINE ON THE 28S RRNA.

CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

DR EMBL; S70176; AAB31048.1; --

DR PIR; J05032; J05032.

DR HSSP; P09989; IMRJ.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.

DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.

DR GO; GO:0015070; F:toxin activity; IEA.

DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.

DR InterPro; IPR001574; RIP.

DR Pfam; PF00161; RIP; 1.

DR PRINTS; PR00396; SHIGARICIN.

DR PROSITE; PS00275; SHIGA_RICIN; 1.

KW Hydrolase; Toxin.

SQ SEQUENCE 289 AA; 31650 MW; 286AC14D48BCA175 CRC64;

Query Match 98.3%; Score 1408; DB 10; Length 289;

Best Local Similarity 97.9%; Pred. No. 2.6e-103;

Matches 283; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

Db 1 MIRFLVLSLLILTLFTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60

QY 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNGASATEAAKYVFK 120

Db 61 RSLPGSQRYALHLNTYADETISVAIDVTNVYIMGYRAGDTSYFFNEASATEAAKYVFK 120

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180

Db 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAASALMWLIQS 180

QY 181 TSEAAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQGFSPVVLIN 240

Db 181 TSEAAARYKFTEQOIGKRVKDTFLPSLAISLENSWSALSQIOIASTNNQGFSPVVLIN 240

QY 241 AQNRQVTITNDAGVVTNSIALLLNNRNNMAAMDDDDVPMTQSFCCGSYAI 289

Db 241 AQNRQVTITNDAGVVTNSIALLLNNRNNMAAMDDDDVPMTQSFCCGSYAI 289

RESULT 4

Q41611 PRELIMINARY; PRT; 270 AA.

ID Q41611

AC Q41611

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Trichosanthes (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
RN NCBI_TaxID=3677;
RP SEQUENCE FROM N.A.
RA Bao Y., Chu R., Han J., Zhang H., Pan N., Gu X., Chen Z.;
RT "Cloning and sequencing of trichosanthin gene and its expression in
RT Escherichia coli and tobacco plant.";
RL Sci. China B. 36:669-676 (0).
RN [2].
RP SEQUENCE FROM N.A.
RA Xu L.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC EMBL; U25675; AAA70036.1; -.
DR HSP; P09989; IMLJ.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.
FT NON_TER 270 270
SQ SEQUENCE 270 AA; 29993 MW; 3D73FB461EA8B8D4 CRC64;
Query Match 88.1%; Score 1262; DB 10; Length 270;
Best Local Similarity 94.8%; Pred. No. 8.4e-92;
Matches 256; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSLPGSQRVYALIHNTVYADETISVAIDVTVNYIMGYRAGDTSYFFNGASATEAAKYVFK 120
DB 61 RSTLPGCERYALIHNTVYADETISVAIDVTVNYIMGYRAGDTFFYFNEASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMVLIQS 180
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMVLIQS 180
QY 181 TSEAAKYKFIEQQIGKRVKDTFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240
DB 181 TSEAAKYKFIEQQIGKRVKDTFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240
QY 241 AQNRVTITNDAGVTVTSNIALLRNNMA 270
DB 241 AQNRVTITNDAGVTVTSNIALLRNNMA 270
RESULT 5
Q8LPV7 ID Q8LPV7 PRELIMINARY; PRT; 270 AA.
AC Q8LPV7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Trichosanthin precursor (Fragment).
GN TCS.
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
RN SEQUENCE FROM N.A.
RP Yuan H., Wang L., Wang Y., Liu T., An C., Chen Z.;
RT "Trichosanthes kirilowii trichosanthin precursor (TCS) gene.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC EMBL; AY082348; AAM22782.1; -.
DR PIR; JCS032; JCS032.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 23
FT CHAIN 24 >270
FT NON_TER 270 270
SQ SEQUENCE 270 AA; 29683 MW; 531713B754F9B769 CRC64;
Query Match 87.5%; Score 1253; DB 10; Length 270;
Best Local Similarity 94.8%; Pred. No. 4.3e-91;
Matches 256; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
DB 1 MIRFLVLSLLILTLFTTTPAVEGDVSFRLSGATSSSYGVFISNLRKALPNERKLYDIPLL 60
QY 61 RSLPGSQRVYALIHNTVYADETISVAIDVTVNYIMGYRAGDTSYFFNGASATEAAKYVFK 120
DB 61 RSLSGSQRVYALIHNTVYADETISVAIDVTVNYIMGYRAGDISYFFNEASATEAAKYVFK 120
QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMVLIQS 180
DB 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSAALMVLIQS 180
QY 181 TSEAAKYKFIEQQIGKRVKDTFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240
DB 181 TSEAAKYKFIEQQIGKRVKDTFLPSLAISLNSWSALSQIQIASTNNGQFESPVLIN 240
QY 241 AQNRVTITNDAGVTVTSNIALLRNNMA 270
DB 241 AQNRVTITNDAGVTVTSNIALLRNNMA 270
RESULT 6
Q9LRE3 ID Q9LRE3 PRELIMINARY; PRT; 247 AA.
AC Q9LRE3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Trichobakin (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).
GN TBK.
OS Trichosanthes sp. Bac Kan 8-98.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=118182;
RN [1].
RP SEQUENCE FROM N.A.
RA Phan Van Chi, Nguyen Thuy Ha, Le Tran Binh;
RT "Genomic DNA Clone for mature typ-1 ribosome-inactivating protein from
RT Trichosanthes sp. sample 01 Bac Kan 8-98 Vien CNHS (Hanoi).";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE

CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; AB039324; BAA92530.1; -.
DR HSP; P09989; 1MRJ.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.
FT NON TER 247 247
SQ SEQUENCE 247 AA; 27199 MW; 89811AC32892F03F CRC64;

Query Match 82.4%; Score 1180; DB 10; Length 247;
Best Local Similarity 96.8%; Pred. No. 2.3e-85;
Matches 239; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 24 DVSRFSGATSSGVFISNLRKALPNERKLYDIPLPSSLPQSQRYALIHLYNVADETI 83
DB 1 DVSRFSGATSSGVFISNLRKALPYERKLYDIPLPSSLPQSQRYALIHLYNVADETI 60

QY 84 SVADVTNVYMGYRAGDTSYFFNGASATEAAKYFKDAMRKVTLPYSGNYERLQTAAGK 143
DB 61 SVADVTNVYMGYRAGDTSYFFNGASATEAAKYFKDAMRKVTLPYSGNYERLQTAAGK 120

QY 144 IRENPLGLPALDSAITTLFFYNANSAASALMVLQISTSEARVYKFIQQIGKRVYDKTFL 203
DB 121 IRENPLGLPALDSAITTLFFYNANSAASALMVLQISTSEARVYKFIQQIGKRVYDKTFL 180

QY 204 PSLAIIISLNSWSALSQKIQIASTNNGQFESPVLINAGNQRVTTINVDAGVVTNSIAL 263
DB 181 PSLAIIISLNSWSALSQKIQIASTNNGQFETPVLINAGNQRVTTINVDAGVVTNSIAL 240

QY 264 LNRNNA 270
DB 241 LNRNNA 247

RESULT 7
Q00980 PRELIMINARY; PRT; 278 AA.
AC Q00980;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE B-luffin (EC 3.2.2.22) (rRNA N-glycosidase).
OS Luffa cylindrica (smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.
OX NCBI_TaxID=3670;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=92288316; PubMed=1600156;
RX Kataoka J., Habuka N., Miyano M., Masuta C., Koiwai A.;
RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
inactivating protein from Luffa cylindrica.";
RL Plant Mol. Biol. 18:1199-1202(1992).
RN [2]
SEQUENCE FROM N.A.
RP MEDLINE=92353400; PubMed=1643290;
RX Kataoka J., Habuka N., Miyano M., Masuta C., Koiwai A.;
RT "Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-
inactivating protein from Luffa cylindrica.";
RL Plant Mol. Biol. 19:887-889(1992).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

DR EMBL; X62372; CAA44230.1; -.
DR PIR; S23519; S23519.
DR HSP; P16094; 1AHC.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.
SQ SEQUENCE 278 AA; 30586 MW; 40C26576EB53F850 CRC64;

Query Match 54.1%; Score 775; DB 10; Length 278;
Best Local Similarity 58.6%; Pred. No. 3e-53;
Matches 163; Conservative 47; Mismatches 64; Indels 4; Gaps 3;

QY 1 MIRELVLSLILFLTPAVG-DVSRFSGATSSGVFISNLRKALPNERKLYDIPL 59
DB 1 MNRFTFLSLILIAFT--VEGANVSFLSGADSKSYKFEITALKALPKEKVNIP 58

QY 60 LRSLPQSQRVALIHLYNVADETTISVALDVTNVYMGYRAGDTSYFFNGASATEAAKYVF 119
DB 59 LLPSASGASRYLMQLSNYDAKAITMADVTNVYMGYRAGDTSYFFNGASATEAAKYVF 118

QY 120 KDAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFFYNANSAASALMVLQ 179
DB 119 KGS-TIVTLPYSGNYERLQNAAGKREKIPGLFRAFDASITSLFHYDSTAAGAPLVIIQ 177

QY 180 STBAARYKFIQQIGKRVYDKTFLPSPALIIISLNSWSALSQKIQIASTNNGQFESPVL 239
DB 178 TTAASREKYIEGQIIEIPKNEVPSAALSLENWSALSQKIQIQAQTNGAFRTPVVII 237

QY 240 NAQNRVITNVDAQVTSNTALLINRNNAAMDDVP 277
DB 238 DNKGORVEIKDVKVNTNNIKLLNKONIAAFDDGIP 275

RESULT 8
Q09FUV7 PRELIMINARY; PRT; 286 AA.
AC Q09FUV7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MAP301 (EC 3.2.2.22) (rRNA N-glycosidase).
OS Momordica charantia (Bitter melon) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3673;
RN [1]
SEQUENCE FROM N.A.
RP Quanhong Y., Rihe P., Aisheng X.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; AF284811; AAG33028.1; -.
DR PIR; B61318; B61318.
DR HSP; P09989; 1MRJ.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.

SQ SEQUENCE 286 AA; 32031 MW; 5B2DF55A41D8F921 CRC64;

Query Match 52.9%; Score 758; DB 10; Length 286;
 Best Local Similarity 59.2%; Pred. No. 6.9e-52;
 Matches 157; Conservative 36; Mismatches 68; Indels 2; Gaps 2;

QY 1 MIRELVLSLLILTLTPPAVEGDVSFRLSGATSSSYGVFISNRKALPNERKLYDIPL 60
 Db 1 MVKCLLSFLIIAIFGVPTAKGVNFDLSTATAKTYTKFIEDFRATLPFSHKYDIPL 60

QY 61 RSLPSQSRVALHNLNYADETISVAIDVTNVIMGYRAGDTSYFENGASATAAKYVFK 120
 Db 61 YSTISDSRRFILLNLTSYAYETISVAIDVTNVVYAYTRDVSYFFK-ESPPEAYNLFK 119

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSASALMVLIQS 180
 Db 120 -GTRKITLPTGNYENLQTAAGKIRENIDGLPALSSAITTLFYNNASAPSALLVLIQT 178

QY 181 TSEARVYKFTQOIGKRVKDTFLPSLAISLNSWALSQKIOIASTNNQFESPVLIN 240
 Db 179 TAEARFKYTERHVAKVATNFKPNLAISLNSWALSQKIOIASTNNQFESPVLIN 240

QY 241 AQNRVTITNDAGVVTNSIALLLN 265
 Db 239 PTGERFQVTNVDSDVVKGNIKLLN 263

RESULT 9

Q41257 ID Q41257 PRELIMINARY; PRT; 286 AA.

AC Q41257; DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE MAP30 (BC 3.2.2.22) (rRNA N-glycosidase).
 GN MAP30.

OS Momordica charantia (Bitter melon) (Balsam pear).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
 OX NCBI_TaxID=3673;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=95394347; PubMed=7665070;
 RA Lee-Huang S., Huang P.L., Chen H.C., Huang P.L., Bourinbalar A.,
 RT Huang H.I., Kung H.F.;
 RT "Anti-HIV and anti-tumor activities of recombinant MAP30 from bitter
 melon.";
 RL Gene 161:151-156 (1995).

CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.

CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

DR EMBL; S79450; AAB35194.2; -.
 DR PIR; B61318; B61318.
 DR PIR; JC4235; JC4235.
 DR PDB; 1D8V; 10-MAY-00.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
 DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
 DR GO; GO:0015070; F:toxin activity; IEA.
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00161; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 DR KW Hydrolase; Toxin.
 DR FT NON TER 1 1
 DR CHAIN 1 2 >264 RIBOSOME INACTIVATING PROTEIN, RIP. TYPE I.

FT NON TER 264 264
 FT SEQUENCE 286 AA; 32018 MW; FA94E5A4A1367C1 CRC64;

Query Match 52.6%; Score 753; DB 10; Length 286;
 Best Local Similarity 58.9%; Pred. No. 1.7e-51;
 Matches 156; Conservative 38; Mismatches 69; Indels 2; Gaps 2;

QY 1 MIRELVLSLLILTLTPPAVEGDVSFRLSGATSSSYGVFISNRKALPNERKLYDIPL 60

QY 61 RSLPSQSRVALHNLNYADETISVAIDVTNVIMGYRAGDTSYFENGASATAAKYVFK 120
 Db 61 YSTISDSRRFILLNLTSYAYETISVAIDVTNVVYAYTRDVSYFFK-ESPPEAYNLFK 119

QY 121 DAMRKVTLPYSGNYERLQTAAGKIRENIPGLPALDSAITTLFYNNANSASALMVLIQS 180
 Db 120 -GTRKITLPTGNYENLQTAAGKIRENIDGLPALSSAITTLFYNNASAPSALLVLIQT 178

QY 181 TSEARVYKFTQOIGKRVKDTFLPSLAISLNSWALSQKIOIASTNNQFESPVLIN 240
 Db 179 TAEARFKYTERHVAKVATNFKPNLAISLNSWALSQKIOIASTNNQFESPVLIN 240

QY 241 AQNRVTITNDAGVVTNSIALLLN 265
 Db 239 PTGERFQVTNVDSDVVKGNIKLLN 263

RESULT 10

Q9FSH2 ID Q9FSH2 PRELIMINARY; PRT; 264 AA.

AC Q9FSH2; DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Ribosome inactivating protein, RIP, type I (EC 3.2.2.22) (rRNA
 DE N-glycosidase) (Fragment).
 GN RIP.

OS Momordica charantia (Bitter melon) (Balsam pear).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
 OX NCBI_TaxID=3673;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen Huy H., Nghiem Ngoc M., Dao Huy P., Le Tran B., Nong Van H.;
 RT "Expression of a RIP gene from Momordica charantia in E. coli.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.

CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

DR EMBL; AJ294541; CAC08217.1; -.
 DR PIR; B61318; B61318.
 DR HSP; P09989; IMR3.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
 DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
 DR GO; GO:0015070; F:toxin activity; IEA.
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00161; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 DR KW Hydrolase; Toxin.
 DR FT NON TER 1 1
 DR CHAIN 1 2 >264 RIBOSOME INACTIVATING PROTEIN, RIP. TYPE I.

FT NON TER 264 264
 FT SEQUENCE 264 AA; 29760 MW; AD9FCC032972718D CRC64;

Query Match 48.8%; Score 699; DB 10; Length 264;
 Best Local Similarity 60.7%; Pred. No. 2.9e-47;
 Matches 147; Conservative 31; Mismatches 62; Indels 2; Gaps 2;

QY 24 DVSPRLSGATSSSYGVFISNRKALPNERKLYDIPLRSLPSQSRVALHNLNYADETI 83
 Db 2 DVNFDLSTATAKTYTKFIEDFRATLPFSHKYDIPLLSYDTSISDSRRFILLNLTSYAYETI 61

QY 84 SVAIDVTNVIMGYRAGDTSYFENGASATAAKYVFKDAMRKVTLPYSGNYERLQTAAGK 143
 Db 62 SVAIDVTNVVYAYTRDVSYFFK-ESPPEAYNLFK-GTRKITLPTGNYENLQTAAGK 119

144 IRENIPGLPALSATITLFFYNANSASALMVLIOSTSEARXKFFIEQQIGKVDKTEL 203
DB 120 IRENIELGLPALSAITLFFYNQASPSALLVLIQTAEARFKYIERHVKAVYATNFK 179
QY 204 PSLAIISLNSWSALSQIOIASTNNGQFSPVVLINAQNRVITINVDAGVTSNIAL 263
DB 180 PNLAIISLNSWSALSQIOIPLAQNQAKFRPNVDLIKPTGERFQVTVNVDSDVVGNIKLL 239
QY 264 LN 265
DB 240 LN 241
RESULT 11
Q8SAD7 PRELIMINARY; PRT; 136 AA.
AC Q8SAD7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Type 1 ribosome-inactivating protein (EC 3.2.2.22) (rRNA
N-glycosidase) (Fragment).
OS Cucurbita moschata (Cushaw squash) (Winter crookneck squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Cucurbitales; Cucurbitaceae; Cucurbita.
OC NCBI_TaxID=3662;
QX 11
RP SEQUENCE FROM N.A.
RA Lin Y., Wu Z.J., Lin Q.Y., Xie L.H.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC -!- SIMILARITY: BELONGS TO THE 28S RRNA.
DR EMBL; AF462349; AAL67855.2;
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.
DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS00231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Glycosidase.
FT NON_TER 136 136
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 15195 MW; F728C9E80B176D61 CRC64;
Query Match 32.8%; Score 469.5; DB 10; Length 136;
Best Local Similarity 67.9%; Pred. No. 1.6e-29;
Matches 93; Conservative 21; Mismatches 22; Indels 1; Gaps 1;
QY 89 VTNVYIMVYAGTSTYFFNGASATEAAKYVFKDAMRKVTLPYSGNYERLQTAAGKIRENI 148
DB 1 VTNVYGVGLVNSTSTFFNEDSQAQSKYVFKDSTR-ITLPYSGNYEKLQAAGKREKI 59
QY 149 PLGLPALDSATITLFFYNANSASALMVLIOSTSEARXKFFIEQQIGKVDKTELPSLAI 208
DB 60 PLGLPALDNTITLFFHYDSTAAAAFTVWIQCTAEASRYRYIEAQMIKRISKDDVPSLAI 119
QY 209 ISLNSWSALSQIOIA 225
DB 120 ISLNSWSALSQIOIFA 136
RESULT 12
Q8GT32 PRELIMINARY; PRT; 563 AA.
AC Q8GT32;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
QY 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Type 2 ribosome-inactivating protein nigrin 1 precursor
(EC 3.2.2.22).
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Dipsacales; Adoxaceae; Sambucus.
OC NCBI_TaxID=4202;
QX 11
RP SEQUENCE FROM N.A.
RA TISSUE=Leaf;
RC Girbes T., Arias P.J., Antolin P.;
RL "Characterization and molecular cloning of Nigrin 1, a type two
ribosome-inactivating protein from leaves of elder (Sambucus nigra).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF249280; AAN86130.1;
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0017149; F:protein biosynthesis inhibitor activity; IEA.
DR GO; GO:0030598; F:RNA N-glycosylase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR InterPro; IPR008997; RicinB-like.
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS00231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Glycosidase.
QX SEQUENCE 563 AA; 62173 MW; 0EB236421FC5E04F CRC64;
Query Match 30.6%; Score 437.5; DB 10; Length 563;
Best Local Similarity 36.2%; Pred. No. 4.2e-26;
Matches 102; Conservative 66; Mismatches 101; Indels 13; Gaps 6;
QY 6 VLSLILITLITTPAVEG---DVSFRLSGATSSYGVFISNLKAL-PNERKLYDIPLL 60
DB 8 MLFYFIVLVAICSVGIQGDYPSVFNLDGAKSATVRDLSNLKRTKATGTVEVNGLPVL 67
QY 61 R--SSLPGSQRYALHILTYADETISVAIDVTNVIYMGVAGDTSYFFNGASATEAAKYV 118
DB 68 RRESEVQVKSREVLPLTNNGTNTVLAVDVNLVYVAFSGNANSYFFK--DATEVQKSN 125
QY 119 FKDAMRKVTLPYSGNYERLQTAAGKIRENIPLGLPALDSATITLFFYNANSASALMVL 178
DB 126 LFVGTKONTLSFGVNDNLETAANTERESIEGLGSPDLGATISL--YHGDSVARSLVVI 183
QY 179 QSTSEAAKYKFFIEQQIGKVDK--TFPLSLAIISLNSWSALSQIOIASTNNGQFSPV 236
DB 184 QMVSEAAARFRIEQLVRRSLQQATSTFPNASMLSMENNWSMSLEIQAGNNVSPFSGTV 243
QY 237 VLINAQNRVITINVDAGVTSNIALLNENMAAMDDDDVPM 278
DB 244 QLINYDTHRLVDNFEELYKITGAILLFRCSPPSNDNAIRM 285
RESULT 13
Q945S2 PRELIMINARY; PRT; 563 AA.
AC Q945S2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA
N-glycosidase).
GN AVL.
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Dipsacales; Adoxaceae; Sambucus.

OX NCBI_TaxID=4202;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Van Damme E.J.M.;
 RT "Characterization and cloning of lectins and ribosome-inactivating
 proteins from *Sambucus nigra* leaves.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 DR EMBL; AF409135; AAL04123.1; -
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0017449; F:protein biosynthesis inhibitor activity; IEA.
 DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
 DR GO; GO:0017148; F:toxin activity; IEA.
 DR GO; GO:0015070; F:negative regulation of protein biosynthesis; IEA.
 DR InterPro; IPR008997; RicinB like.
 DR InterPro; IPR000772; RicinB_lectin.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00652; RicinB_lectin; 6.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 KW Hydrolase; Toxin.
 SQ SEQUENCE 563 AA; 62242 MW; 07F7CBEDCF33BF10 CRC64;
 Query Match 30.3%; Score 433.5; DB 10; Length 563;
 Best Local Similarity 35.9%; Pred. No. 8.8e-26;
 Matches 103; Conservative 66; Mismatches 105; Indels 13; Gaps 6;
 QY 1 MIRFLVLSLLILTLFTTTPAVEG----DVSFRISGATSSSYGVFISNLRKAL-PNERKLY 55
 Db 3 VIAAAMLYFYVVLVLAICSVGIQIDYPSVSNFLDGAKSATYRDFLNLRKTIVATGYEVN 62
 QY 56 DIPLLR--SSLPQSRYALIHLYNVADETSIVADVTNVTYIMGYRAGDTSYFENGASATE 113
 Db 63 GLPVLRESEVQVKRFVLPLTNNGTNTVLAVDVNLYVVAFSNANSYFFK--DATE 120
 QY 114 AAKYVFKDAMRKVTLPYSGNYERLOTAAGKIRENIPGLPALDSAITTLFFYNANSAASA 173
 Db 121 VQKSNLFVCTKNTLSFTNGYDNLETAAGTRESIELGSPDLGATSL--YHGSVARS 178
 QY 174 LMVLIQSTSEAAARYKFIQOIGKRVDK--TFLPSLAISLNSWSALSQIOIASTNNQO 231
 Db 179 LLVVIQMVSEAAARFYIEQEVRSRLOQATSFTPNALMLSMENNWSMSLEIQAGNNVSP 238
 QY 232 FESPVLINAQNRVTITNDAGVVTSTNIALLRNNMAMDDDDVPM 278
 Db 239 FPGTVOLLNVDYHTRLVDFNFEELYKITGIAILLFRCSPPSNDNAIRM 285
 RESULT 14
 ID 004367 PRELIMINARY; PRT; 563 AA.
 AC 004367;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Ribosome inactivating protein precursor (BC 3.2.2.22) (rRNA
 N-glycosidase).
 DE *Sambucus nigra* (European elder).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; ascerids;
 OC campanulids; Dipsacales; Adoxaceae; *Sambucus*.
 OX NCBI_TaxID=4202;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98112023; PubMed=9450339;
 RA Van Damme E.J., Roy S., Barre A., Ronge P., Van Leuven F.,
 RA Peumans W.J.;

"The major elderberry (*Sambucus nigra*) fruit protein is a lectin
 derived from a truncated type 2 ribosome-inactivating protein.";
 RL Plant J. 12:1251-1260(1997).
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 DR EMBL; U76524; AAC15886.1; -
 DR HSP; P02879; 2AAL
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0017449; F:protein biosynthesis inhibitor activity; IEA.
 DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
 DR GO; GO:0015070; F:toxin activity; IEA.
 DR GO; GO:0017148; F:negative regulation of protein biosynthesis; IEA.
 DR InterPro; IPR008997; RicinB like.
 DR InterPro; IPR000772; RicinB_lectin.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00652; RicinB_lectin; 6.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 KW Hydrolase; Signal; Toxin.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 297 RIBOSOME INACTIVATING PROTEIN, A CHAIN.
 FT CHAIN 298 563 RIBOSOME INACTIVATING PROTEIN, B CHAIN.
 SQ SEQUENCE 563 AA; 62336 MW; 3ED2B6C08E796205 CRC64;
 Query Match 30.1%; Score 430.5; DB 10; Length 563;
 Best Local Similarity 37.1%; Pred. No. 1.5e-25;
 Matches 102; Conservative 62; Mismatches 98; Indels 13; Gaps 6;
 QY 1 MIRFLVLSLLILTLFTTTPAVEG----DVSFRISGATSSSYGVFISNLRKAL-PNERKLY 55
 Db 3 VIAAAMLYFYVVLVLAICSVGIQIDYPSVSNFLDGAKSATYRDFLNLRKTIVATGYEVN 62
 QY 56 DIPLLR--SSLPQSRYALIHLYNVADETSIVADVTNVTYIMGYRAGDTSYFENGASATE 113
 Db 63 GLPVLRESEVQVKRFVLPLTNNGTNTVLAVDVNLYVVAFSNANSYFFK--DATE 120
 QY 114 AAKYVFKDAMRKVTLPYSGNYERLOTAAGKIRENIPGLPALDSAITTLFFYNANSAASA 173
 Db 121 LQKSNLFVCTKNTLSFTNGYDNLETAAGTRESIELGSPDLGATSLY--DESVARS 178
 QY 174 LMVLIQSTSEAAARYKFIQOIGKRVDKT--FLPSLAISLNSWSALSQIOIASTNNQO 231
 Db 179 LLVVIQMVSEAAARFYIEQEVRSRLOQATSFTPNALMLSMENNWSMSLEVOQSGDNVSP 238
 QY 232 FESPVLINAQNRVTITNDAGVVTSTNIALLRN 266
 Db 239 FGTGTVOLLNVDYHTRLVDFNFEELYKITGIAILLFR 273
 RESULT 15
 ID 041174 PRELIMINARY; PRT; 541 AA.
 AC 041174;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Proridin A chain (EC 3.2.2.22) (rRNA N-glycosidase)
 DE (Fragment).
 OS *Ricinus communis* (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaceae;
 OC *Ricinus*.
 OX NCBI_TaxID=3988;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92338377; PubMed=1633311;
 RA Roberts L.M., Tregear J.W., Lord J.M.;
 RA "Molecular cloning of ricin.";

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: September 14, 2004, 18:55:11 ; Search time 3322 Seconds
(without alignments)
3770.661 Million cell updates/sec
Title: US-09-905-247-1
Perfect score: 1432
Sequence: 1 MRFVLVSLILTLFTTPA.....AAMDVDPMTQSGFGCSYAL 289

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cn2 1/USPTO spool_p/US09905247/runat_10092004_103933_4713/app_query.fasta_1.455
-DB=GenEmbl -QMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09905247 @CGN 1 1 5600 @runat_10092004_103933_4713 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl :
1: gb_ba :
2: gb_htg :
3: gb_in :
4: gb_om :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_rts :
11: gb_scs :
12: gb_sy :
13: gb_un :
14: gb_vi :
15: em_ba :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_or :
22: em_ov :
23: em_pat :
24: em_ph :
25: em_pl :
26: em_ro :
27: em_sts :
28: em_un :

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1422	99.3	1010	8	TRHTCSA	M34858 T.kirillowii
2	1419	99.1	1809	8	AY082349	AY082349 Trichosan
3	1418	99.0	1799	8	AF367252	AF367252 Trichosan
4	1413	98.7	886	6	AR032228	AR032228 Sequence 3
5	1413	98.7	886	6	I32881	I32881 Sequence 3
6	1413	98.7	886	6	AR228702	AR228702 Sequence
7	1413	98.7	886	6	BD011600	BD011600 Recombina
8	1413	98.7	886	6	BD165859	BD165859 Transgeni
9	1413	98.7	1227	8	TRHTCS	J05434 T.kirillowii
10	1408	98.3	1134	8	S70176	S70176 trichosanth
11	1377	96.2	1194	8	AB000666	AB000666 Trichosan
12	1262	88.1	810	8	TKU25675	U25675 Trichosanth
13	1253	87.5	1433	8	AY082348	AY082348 Trichosan
14	1180	82.4	741	8	AB039324	AB039324 Trichosan
15	1164.5	81.3	1084	8	BYNBD1	BYNBD1 Bryonia dio
16	1164.5	81.3	1094	6	I24020	I24020 Sequence 1
17	1074	75.0	1499	6	AR117137	AR117137 Sequence 1
18	1064	74.3	758	6	AR117133	AR117133 Sequence
19	859.5	60.0	1044	8	MCAMC	X57682 M.charantia
20	802.5	56.0	1026	6	E03988	E03988 DNA encodin
21	802.5	56.0	1026	8	LCALUF	X62371 L.cylindric
22	775	54.1	914	6	E03989	E03989 DNA encodin
23	775	54.1	914	8	LCBLUF	X62372 L.cylindric
24	766.5	53.5	1012	6	E07538	E07538 DNA sequenc
25	758	52.9	861	8	AF284811	AF284811 Momordica
26	756	52.8	914	8	MBMOMORII	Z12175 M.balsamina
27	753	52.6	919	8	S79450	S79450 Map30=anti-
28	736	51.4	1042	8	AB045560	AB045560 Cucumis f
29	728	50.8	913	6	AR003723	AR003723 Sequence
30	728	50.8	913	6	AR010059	AR010059 Sequence
31	728	50.8	913	6	AR055265	AR055265 Sequence
32	728	50.8	913	6	AR141194	AR141194 Sequence
33	728	50.8	913	6	AR141431	AR141431 Sequence
34	728	50.8	913	6	I11903	I11903 Sequence 13
35	728	50.8	913	6	I40476	I40476 Sequence 13
36	728	50.8	913	6	AR367960	AR367960 Sequence
37	728	50.8	913	6	AR430941	AR430941 Sequence
38	699	48.8	795	8	MCH294541	AJ294541 Momordica
39	685.5	47.9	1039	8	AF055086	AF055086 Trichosan
40	537.5	37.5	962	6	I34238	I34238 Sequence 14
41	469.5	32.8	408	8	AF462349	AF462349 Cucurbita
42	437.5	30.6	1906	8	AF249280	AF249280 Sambucus
43	433.5	30.3	1692	8	AF409135	AF409135 Sambucus
44	432.5	30.2	1879	8	SNU41299	U41299 Sambucus ni
45	430.5	30.1	2079	8	SNU76524	U76524 Sambucus ni

ALIGNMENTS

Db	1420	ACCTTTGTTTACTACACGCCAATTCGTGCTGCGCACTTATGGTACTCATTCAGTCG	1479
QY	181	ThrSerGluAlaAlaArgTVrLysPheIleGluGlnIleGlyLysArgValAspLys	200
Db	1480	ACGTCTGAGCTCGGAGGTATAAATTTATTGACAAATAATGGGAAGCGGTGTGACAAA	1539
QY	201	ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTrpSerAlaLeuSerLys	220
Db	1540	ACCTTCTCAACCAAGTTTAGCAATTTATAGTTTGGAAATAGTTGGTCTGCTCTCTCCAAG	1599
QY	221	GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn	240
Db	1600	CAAAATTACAGATACGAGTACTAATAATATGACAGTTTGAAGTCCTGTGTGCTTATAAT	1659
QY	241	AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle	260
Db	1660	GCTCAAAACCAACGAGTCAGATAACCAATGTGTGCTGGAGTTGTAACCTCCAAATC	1719
QY	261	AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln	280
Db	1720	GCCTTCTGCTGAATAGAACCAATATGGCAGCCATGGATGACGATGTTCTATGACACAG	1779
QY	281	SerPheGlyCysGlySerTyrAlaLeu	289
Db	1780	AGCTTTGGATGTGGAAGTTATGCTATT	1806
RESULT 3			
AF367252			
LOCUS			
DEFINITION	AF367252	1799 bp DNA linear PLN 14-MAY-2002	
ACCESSION	AF367252	Trichosanthes kirilowii trichosanthin precursor (TCS) gene,	
VERSION	AF367252	complete cds.	
KEYWORDS	AF367252.1	GI:14030568	
SOURCE			
ORGANISM		Trichosanthes kirilowii (Mongolian snake-gourd)	
REFERENCE			
AUTHORS		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
TITLE		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
JOURNAL		rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.	
FEATURES		1 (bases 1 to 1799)	
source		Yuan,H., Wang,L., Wang,Y., An,C. and Chen,Z.	
		Direct Submission	
		Submitted (31-MAR-2001) National Laboratory of Protein Engineering	
		and Plant Genetic Engineering, College of Life Sciences, Peking	
		University, Beijing 100871, China	
		Location/Qualifiers	
		1..1799	
		/organism="Trichosanthes kirilowii"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:3677"	
		/clone="type2"	
gene		1..>1799	
		/gene="TCS"	
promoter		1..904	
		/gene="TCS"	
misc_feature		903	
		/gene="TCS"	
		/note="alternative weaker transcriptional start site"	
mRNA		905..>1799	
		/gene="TCS"	
5'UTR		/product="trichosanthin precursor"	
		905..929	
		/gene="TCS"	
misc_feature		906	
		/gene="TCS"	
		/note="alternative weaker transcription start site"	
CDS		930..1799	
		/gene="TCS"	
		/function="inactivates eukaryotic 60S ribosomal subunits"	
		/note="preproprotein"	
		/codon_start=1	
		/product="trichosanthin precursor"	

AR032228	LOCUS	DEFINITION	Sequence 3 from patent US 5866785.	886 bp	DNA	linear	PAT 29-SEP-1999
AR032228	ACCESSION	AR032228	AR032228.1	GI:5946517			
AR032228	VERSION	AR032228.1	GI:5946517				
AR032228	KEYWORDS	Unknown.					
AR032228	SOURCE	Unknown.					
AR032228	ORGANISM	Unknown.					
AR032228	REFERENCE	1 (bases 1 to 886)					
AR032228	AUTHORS	Donson, J., Dawson, W.O., Grantham, G.L., Turpen, T.H., Turpen, A., Myers, Garger, S.J., and Grill, L.K.					
AR032228	TITLE	Recombinant plant viral nucleic acids					
AR032228	JOURNAL	Patent: US 5866785-A 3 02-FEB-1999;					
AR032228	FEATURES	Location/Qualifiers					
AR032228	source	1..886					
AR032228		/organism="unknown"					
AR032228		/mol_type="unassigned DNA"					
AR032228	ORIGIN						
AR032228	Alignment Scores:						
AR032228	Pred. No.:	2.85e-119	Length:	886			
AR032228	Score:	1413.00	Matches:	285			
AR032228	Percent Similarity:	99.31%	Conservative:	2			
AR032228	Best Local Similarity:	98.62%	Mismatches:	0			
AR032228	Query Match:	98.67%	Indels:	0			
AR032228	DB:	6	Gaps:	0			
AR032228	US-09-905-247-1 (1-289) x AR032228 (1-886)						
AR032228	Qy	1 MettleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrThrProAla	20				
AR032228	Db	8 ATGATCAGATTCTTAGTCCTCTCTTTGGTAATTCACCCCTTCTCTAACAACTCTGCT	67				
AR032228	Qy	21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe	40				
AR032228	Db	68 GTGGAGGGCGATGTAGCTTCGGTTTATCAGGTGCAACAAGCAGTTCCTATGGAGTTTTC	127				
AR032228	Qy	41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu	60				
AR032228	Db	128 ATTTCAAATCTGAGAAAAGCTCTTCCAATGAAGGAACCTGACGATATCCCTCTGTGA	187				
AR032228	Qy	61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp	80				
AR032228	Db	188 CGTTCTCTCTTCCAGGTTCTCAAGCTACGATGTATCCATCTCACAAATACGCCGAT	247				
AR032228	Qy	81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly	100				
AR032228	Db	248 GAAACCAATTCAGTGGCCATAGACGTACGAAACGCTATATTATGGGATATCGCGCTGGC	307				
AR032228	Qy	101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys	120				
AR032228	Db	308 GATACATCTATTTTTTCAACGAGGCTTCTGCAACAGAAAGCTGCAAAATATGTATTCAA	367				
AR032228	Qy	121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla	140				
AR032228	Db	368 GACGCTATCGGAAAAGTTACGTTCTCATATTCTGGCAATTACGAAAGGCTTCAAACCTGCT	427				

```
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 428 GCGGGCAAAATAAGGAAATAATTCGCTGGACTCCAGCTTGGACATGGCATTAAC 487
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 488 ACTTTGTTTTACTACACGCAATTCCTGCTGCGCTCGCACTTATGGTACTCAITTCAGTCG 547
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
Db 548 ACGTCTGAGGCTGCGAGGTATAAATTTATGACAAACAAATTCGGAAGCGCGTTGACAAA 607
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
Db 608 ACCCTTCTACCAAGTTAGCAATATAAGTTTGAAAAATAGTTGGTCTCTCTCCAAG 667
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 668 CAAATTCAGATACGAGTACTAATAATGACACAGTTTGAAACTCCTGTGTGCTTATAAAT 727
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 728 GCTCAAAACCAACGATCATGATAACCAATGTTGATGCTGGAGTTGTAACTTCCAACATC 787
QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db 788 GCGTTGCTGCTGAATCGAAACAAATATGGCAGCCATGGATGACGATGTTCTATGACACAG 847
QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
Db 848 AGCTTTGGATGTGGAAGTTATGCTATT 874

RESULT 5
LOCUS I32881 886 bp DNA linear PAT 06-FEB-1997
DEFINITION Sequence 3 from patent US 5589367.
ACCESSION I32881
VERSION I32881.1 GI:1823672
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Donson,J., Dawson,W.O., Grantham,G.L., Turpen,T.H., Turpen,A.M.,
Garger,S.J. and Grill,L.K.
TITLE Recombinant plant viral nucleic acids
JOURNAL Patent: US 5589367-A 3 31-DEC-1996;
FEATURES Location/Qualifiers
source
1..886
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 2.85e-119 Length: 886
Score: 1413.00 Matches: 285
Percent Similarity: 99.31% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 2
Query Match: 98.67% Indels: 0
DB: 6 Gaps: 0

US-09-905-247-1 (1-289) x I32881 (1-886)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrThrProAla 20
Db 8 ATGATCAGATCTTAGTCTCTCTTGTCTTAATCTCAACCTTCTTCAACAACTCCTGCT 67
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 68 GTGAGGGCGGATGTTAGTCTCGTTTATCAGGTGCAACAGCAGTTCCTATGGAGTTTTC 127
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
```

```
Db 128 ATTTCAAATCTGAAAAAGCTCTTCCAAATGAAGAAAGAACTGTACGATATCCCTCTCTTA 187
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
Db 188 CGTTCTCTCTTCCAGGTTCTCAACGCTAGCATTGATCCATCTCACAATTTAGCCGAT 247
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db 248 GAAACCATTTTCAGTGGCCATAGACGTACACAACTCTATATTATGGGATATCGCGTGGC 307
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Db 308 GATACATCTCTATTTTTCACGAGGCTTCTGCACAGAAAGCTCAAAATATGATTACAA 367
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 368 GAGCTATTCGAAAAGTTACGCTTCCATATTCCTGGCAATTCGAAAGGCTTCAACTGCT 427
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 428 GCGGGCAAAATAAGGAAATAATTCGCTTGGACTCCAGCTTGGACAGTGCATTACC 487
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 488 ACTTTGTTTTACTACACGCAATTCCTGCTGCGCACTTATGGTACTCAITTCAGTCG 547
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
Db 548 ACGTCTGAGGCTGCGAGGTATAAATTTATGAGCAACAAATTTGGGAAGCGCGTTGACAA 607
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
Db 608 ACCCTTCTACCAAGTTAGCAATATAAGTTTGAAAAATAGTTGGTCTCTCTCTCCAAG 667
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 668 CAAATTCAGATACGAGTACTAATAATGACACAGTTTGAAACTCCTGTGTGCTTATAAAT 727
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 728 GCTCAAAACCAACGATCATGATAACCAATGTTGATGCTGGAGTTGTAACTTCCAACATC 787
QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db 788 GCGTTGCTGCTGAATCGAAACAAATATGGCAGCCATGGATGACGATGTTCTATGACACAG 847
QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
Db 848 AGCTTTGGATGTGGAAGTTATGCTATT 874

RESULT 6
LOCUS AR228702 886 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 3 from patent US 6448046.
ACCESSION AR228702
VERSION AR228702.1 GI:27267789
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Donson,J., Dawson,W.O., Grantham,G.L., Turpen,T.H., Turpen,A.M.,
Garger,S.J. and Grill,L.K.
TITLE Recombinant animal viral nucleic acids
JOURNAL Patent: US 6448046-A 3 10-SEP-2002;
FEATURES Location/Qualifiers
source
1..886
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
```

RESULT 7

```

Db      488 ACTTTGTTTTACTACACGCAATTCCTGCTGCGCACTTATGCTACTCATTCAGTCG 547
QY      181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200
Db      548 ACCTCTGAGCTCGAGGTATAATTTATTTAGACACAAATGGGAAGCGGTGACAA 607
QY      201 ThrPheLeuProSerLeuLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
Db      608 ACCTTCCTACCAAGTTTAGCAATTATAAGTTTGGAAAATAGTTGGTCTCTCTCAAG 667
QY      221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db      668 CAAATTGAGATGCGAGTACTATAATAATGGACAGCTTTGAAACTCCTGCTGCTTATAAT 727
QY      241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db      728 GCTCAAAACCAACGAGTCATGATCAACCAATGTTGCTGGAGTTGTAACCTCCACATC 787
QY      261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db      788 GCGTTGCTGCTGAATCGAAACAATATGGCAGCATGATGACGATGTTCTATGACACAG 847
QY      281 SerPheGlyCysGlySerTyrAlaLeu 289
Db      848 AGCTTTGGATGTGGAAGTTATGCTATT 874

RESULT 8
LOCUS      BD165859                      886 bp      DNA      linear      PAT 17-JAN-2003
DEFINITION Transgenic plant virus nucleic acid.
ACCESSION  BD165859
VERSION    BD165859.1 GI:27871671
KEYWORDS  JP 2002171976-A/2.
SOURCE    unidentified
ORGANISM  unclassified.
REFERENCE  1 (bases 1 to 886)
AUTHORS   Donson,J., Dawson,W.O., Grantham,G.L., Turpen,T.H., Turpen,A.M.,
          Gerger,S.J. and Grill,L.K.
TITLE     Transgenic plant virus nucleic acid
JOURNAL   Patent: JP 2002171976-A 2 18-JUN-2002;
          LARGE SCALE BIOLOGY CORP
COMMENT   OS Cucumis L. (chinese cucumber)
          PN JP 2002171976-A/2
          PD 18-JUN-2002
          PR 01-OCT-2001 JP 2001313207
          PI JOHN DONSON, WILLIAM O DAWSON, GEORGE L GRANTHAN, THOMAS H
          TURPEN,
          PI ANN MYERS TURPEN, STEPHEN J GERGER, LAURENCE K GRILL PC
          C12N15/09, C12N15/00
          CC Clone: alpha-trichosanthin
          FH Key Location/Qualifiers
          FT CDS Location/Qualifiers
          source
          1..886
          /organism="unidentified"
          /mol_type="genomic DNA"
          /db_xref="taxon:32644"

ORIGIN

Alignment Scores:
Pred. No.:      2.85e-119      Length:      886
Score:          1413.00      Matches:      285
Percent Similarity: 99.31%      Conservative: 2
Best Local Similarity: 98.62%      Mismatches: 0
Query Match:    98.67%      Indels:      2
DB:             6      Gaps:      0

US-09-905-247-1 (1-289) x BD165859 (1-886)
QY      1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20

```

```

Db      8 ATGATCAGATTCTTAGTCTCTCTTTTCTAATTTCTCACCTCTTCTCTTAAACAATCTCTGCT 67
QY      21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db      68 GTGAGGGCGCATGTTAGCTTCGTTTATCAGGTGCAACAGCAGTTCCTATGAGATTTC 127
QY      41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db      128 ATTTCAAAATCTGAGAAAAGCTCTTCCAAATGAAAGAAACTGTACGATATCCCTCTGTTA 187
QY      61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleIleIleuThrAsnTyrAlaAsp 80
Db      188 CGTTCCTCTCTCCAGTCTTCAACGCTACGCAATGATCCATCTCAAAATACGCCGAT 247
QY      81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db      248 GAAACCAATTCAGTGCCCATAGACGTAAACGAGCTCTATATTATGGGATATCGCGTGGC 307
QY      101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Db      308 GATACATCTCTATTTTTCACGAGGCTTCTGCAACAGAGCTGCAAAATATGATTATCAA 367
QY      121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db      368 GACGCTATGCGAAAAGTTACGCTTTCATATTCTGGCAATTACGAAGGCTTCAAACTGCT 427
QY      141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db      428 GCGGGCAAAATTAAGGGAAAATATTCCGCTTGGACTCCAGCTTTGGACAGTCCATTACC 487
QY      161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db      488 ACTTTGTTTTACTACACGCCAAATTCGTGCTGCGCACTTATGCTACTCATTCAGTCG 547
QY      181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200
Db      548 AGCTCTGAGGCTCGCAGGTATAAATTTATGAGCAACAAATTTGGGAAGCGCTTGACAAA 607
QY      201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
Db      608 ACCTTCCTACCAAGTTTAGCAATTATAAGTTTGGAAAATAGTTGCTGCTCTCTCCAG 667
QY      221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db      668 CAAATTGAGATGCGAGTACTATAATAATGGACAGTTCGAACTCCTCTGCTGCTTATAAT 727
QY      241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db      728 GCTCAAAACCAACGAGTCATGATCAACCAATGTTGATGCTGGAGTTGTAACTCCCAATC 787
QY      261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
Db      788 GCGTTGCTGCTGAATCGAAACAATATGGCAGCCATGATGACGATGTTCTATGACACAG 847
QY      281 SerPheGlyCysGlySerTyrAlaLeu 289
Db      848 AGCTTTGGATGTGGAAGTTATGCTATT 874

RESULT 9
LOCUS      TRHTCS                      1227 bp      DNA      linear      PLN 16-JUL-1993
DEFINITION T.kirilowii alpha-trichosanthin (alpha-TCS) gene, complete cds.
ACCESSION  J05434
VERSION    J05434.1 GI:170534
KEYWORDS  alpha-trichosanthin; ribosome-inactivating protein; type I
          ribosome-inactivating protein.
SOURCE     Trichosanthes kirilowii (Mongolian snake-gourd)
ORGANISM  Trichosanthes kirilowii
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
REFERENCE  1 (bases 1 to 1227)

```

AUTHORS Chow,T.P., Feldman,R.A., Lovett,M. and Piatak,M.
TITLE Isolation and DNA sequence of a gene encoding alpha-trichosanthin,
a type I ribosome-inactivating protein
JOURNAL J. Biol. Chem. 265 (15), 8670-8674 (1990)
MEDLINE 90256790
PUBMED 2341400
COMMENT Original source text: T.kirilowii (strain Maximowicz) leaf DNA,
clone pQ2D.
Draft entry and computer-readable sequence for [1] kindly submitted
by M.Piatak, 15-MAR-1990, for release after publication.
FEATURES
source 1..1227
/organism="Trichosanthus kirilowii"
/mol_type="genomic DNA"
/db_xref="taxon:3677"
25..30
/note="putative"
TATA_signal 168..173
TATA_signal 238..243
TATA_signal 283..288
TATA_signal 340..1209
/function="inactivates eukaryotic 60S ribosomal subunits"
CDS
/codon_start=1
/product="trichosanthin"
/protein_id="AAA34206.1"
/db_xref="GI:170535"
/translation="MIRPLVLSLILFLTPAVGVDSFRLSGATSSSYGVFISNL
RKALPNRKLVDIPLLRSSLPQSRVALIHNTVNADETISVAIDVTNVIIMGYRAGDT
SYFNEASATPAKVKPDAMKVTLPYSGNVERLOTAAGKIRENIPLGLPALDSAIT
ILFYNEASASALMVLQSTSEARVYFIEQIQIGKRVDTFLPSLAIISLENSWAL
SKQIQIASTNNQGFETPVLVNAQNRVITNVDGVVTSNIALLLNNNAAMDDVDV
PMTQSGGGSYAI"
340..408
/note="alpha-trichosanthin signal peptide"
409..1149
/product="alpha-trichosanthin"
ORIGIN
Alignment Scores:
Pred. No.: 4,388-119 Length: 1227
Score: 1413.00 Matches: 285
Percent Similarity: 99.31% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 2
Query Match: 98.67% Indels: 0
DB: 8 Gaps: 0
US-09-905-247-1 (1-289) x TRHTCS (1-1227)
QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuLeuThrLeuPheLeuThrProAla 20
Db 340 ATGATCAGATTCTTAGTCTCTCTTTGCTAAATCTCACCTCTTCTCAACACTCTCGTCT 399
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerGlyValPhe 40
Db 400 GTGGAGGGCGATGTAGCTTCCTTTATCAGGTGCAACAGAGATTCCTATGGAGTTTC 459
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuThrAspLeuProLeuLeu 60
Db 460 ATTTCAAATCTGAGAAAGCTCTTCCAAATGAAAGGAAACTGTACGATATCCCTCTGTTA 519
QY 61 ArgSerSerLeuProGlySerGlnArgGlyAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Db 520 CGTTCCTCTCTCCAGGTCTTCAACGCTACGCAATGATCCATCTCAAAATACGCCGAT 579
QY 81 GluThrIleSerValAlaIleAspValThrAsnValThrIleMetGlyThrArgAlaGly 100
Db 580 GAAACCATTTTCAGTGCCATAGACGTACGAAAGCTCTATATTATGGATATCGCGTGGC 639
QY 101 AspThrSerThrPhePheAsnGlyAlaSerAlaThrGluAlaLysThrValPheLys 120

Db 640 GATACATCTCTATTTTTTCAACGAGGCTTCTGCAACAGAACTGCAAAATATGTATTCAA 699
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 700 GAGCTATGCGAAAGATTACGCTTCCATATCTTGGCAATTTACGAAAGGCTTCAAACTGCT 759
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 760 CGCGGCAAAATAAGGAAATAATTTCCGCTTGGACTCCAGCTTTGGACAGTGCATTACC 819
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 820 ACTTTGTTTACTACAACGCCAATTTCTGCTGCGTCCGCACTTATGCTACTCATTCAGTCG 879
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGlnGlnIleGlyLysArgValAspLys 200
Db 880 ACCTCTGAGGCTGCGAGGTATTAATTTATTGAGCAACAATTTGGGAAGCGGCTTGACAAA 939
QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrSerAlaLeuSerLys 220
Db 940 ACCTTCTCTACCAAGTTTAGCAATTTAAGTTTGGAAATAGTTGGTCTGCTCTCTTCAAG 999
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 1000 CAAATTCTAGTAGGAGTACTAATATATGACAGCTTTGAAACTCTCTGTTGCTTATAAT 1059
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 1060 GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGGAGTTGTAACCTCCAACATC 1119
QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspAspValProMetThrGln 280
Db 1120 CGCTTGTCTGCTGAATCGAACAATATGCGACCACTGGATGACGATGTTTCTTATGACACAG 1179
QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
Db 1180 AGCTTTGGATGTGGAAGTTATGCTATT 1206
RESULT 10
S70176 1134 bp DNA linear PLN 23-SEP-1994
trichosanthin [Trichosanthus kirilowii, Maxim, Genomic, 1134 nt].
S70176 GI:547148
S70176.1
Trichosanthus kirilowii (Mongolian snake-gourd)
Trichosanthus kirilowii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
1 (bases 1 to 1134)
Zheng,H., Wang,B., Shaw,P. and Yeung,H.
Cloning and DNA sequencing of the gene encoding trichosanthin
Yi Chuan Xue Bao 21 (1), 42-51 (1994)
94271613
8003348
PUBMED
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsg 148037] from the original journal article.
This sequence comes from Fig. 4.
FEATURES
source location/Qualifiers
1..1134
/organism="Trichosanthus kirilowii"
/mol_type="genomic DNA"
/db_xref="taxon:3677"
244..1113
/gene="trichosanthin, TCS"
244..1113
/gene="trichosanthin, TCS"
/note="This sequence comes from Fig. 4; TCS"
/codon_start=1
/product="trichosanthin"
/protein_id="AAB31048.1"
/db_xref="GI:547149"

Alignment Scores:				
Pred. No.:	1.12e-118	Length:	1134	
Score:	1408.00	Matches:	283	
Percent Similarity:	99.65%	Conservative:	5	
Best Local Similarity:	97.92%	Mismatches:	1	
Query Match:	98.32%	Indels:	0	
DB:	8	Gaps:	0	
US-09-905-247-1 (1-289) x S70176 (1-1134)				
QY	1	MetileArgPheLeuValLeuSerLeuLeuLeuLeuLeuThrLeuPheLeuThrProAla	20	
Db	244	ATGATCAGATCTTAGTCCTCTCTTGCTAAATCTCACCTCTTCTTCAACACTCTCTGCT	303	
QY	21	ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe	40	
Db	304	GTGGAGGGCGATGTTAGCTTCGTTTATCAGTGCACCAAGCAGTTCCTATGGAGTTTTC	363	
QY	41	IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspLeuLeu	60	
Db	364	ATTTCAAATCTGAGAAAGCTCTTCCAAATCAAGAGAACTGTACGATATCCCTTGTGTA	423	
QY	61	ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuHleHisLeuThrAsnTyrAlaAsp	80	
Db	424	CGTTCAGTCTTCCAGTCTTCAACGCTACGATGGTCCATCTCACAAATACGCCGAT	483	
QY	81	GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly	100	
Db	484	GAAACCATTTTCAGTGGCCATAGACGTAAACGAGCGTCTATATTATGGGATATCGGCTGCG	543	
QY	101	AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaLysTyrValPheLys	120	
Db	544	GATACATCTCTATTTTTCACAGGAGCTTCTGCAACAGAAAGTGCAAAATATGTATTCAA	603	
QY	121	AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla	140	
Db	604	GACGCTATGCGAAAGTTACGCTTCCATNTCTGCGAATACGAAAGGCTTCAAACTGCT	663	
QY	141	AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr	160	
Db	664	CGCGGCAAAATAAGGGAAAAATATTCGCTTGGACTCCAGCTTTGGACAGTGCCATTACC	723	
QY	161	ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer	180	
Db	724	ACTTTGTTTACTACAAACGCCAATTCCTGCTGCGTCGGCCTTATATGTACTCATTCAGTCG	783	
QY	181	ThrSerGluAlaAlaArgTyrLysPheIleGlnGlnIleGlyLysAtqValAspLys	200	
Db	784	ACGTCTGAGGCTCGAGGTATAATTTATTGAGCAACAATTTGGAGACGGCGTTGACAAA	843	
QY	201	ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys	220	
Db	844	ACCTCTCTACCAAGTTTAGCAATATATAAGTTTGGAAAAATAGTTGGTCTGCTCTCCAAAG	903	
QY	221	GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn	240	
Db	904	CAAAATTCAGATACGGAGTACTAATAATATGGACAGTTTGAACACTCTGTTGTCTTATAAT	963	
QY	241	AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle	260	
Db	964	GCTCAAAACCAACGAGTCAAGATAACCAATGTTGATGCTGGAGTTGTAACCTCCAAATC	1023	
QY	261	AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln	280	
Db	1024	GCCTTCCTGCTGAATCGAAACGATATGGCAGCCATGTGATGACGATGTTCTATGACACAG	1083	

Db 367 GTGAGGCGCATGTAGCTTCGGTTTATCAGGTGCAACAGCAGTTCCTATGGAGTTTC 426
 Qy 41 IleserAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
 Db 427 ATTTCAATCTGAGAAAGCTCTTCCATATGAAAGGAACTGTACGATATCCCTCTCTGTA 486
 Qy 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
 Db 487 CGTTCACACTCTCCAGGTTCTCAACGCTACCACTGATCCATCTCACAATAATACGCCAT 546
 Qy 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
 Db 547 GAACACATTTTCAGTGGCCATAGAGCTAGACAGCTCTATGTTATGGATATCGCGTGT 606
 Qy 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
 Db 607 GATACATCTCTATTTTCAACAGGCTTCTGCAACAGAGCTGCAAAATATGTATTCAA 666
 Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluAlaGluThrAla 140
 Db 667 GACGCTAAGCGAAAGATTACGCTTCCATATTCTGGCAATTACGAAGGCTTCAATTGCT 726
 Qy 141 AlaGlyLysIleAArgLysIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
 Db 727 GCAGCAAAATAGGAAATATTCGCTTGGACTTCCAGCTTGGCAGTGGCATTACC 786
 Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
 Db 787 ACTTGTGTTTACTACAACGCCAATCTGCTGCTGCGCACTTATGCTACTCAATTCAGTCG 846
 Qy 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200
 Db 847 ACGTCTGAGGCTGCGAGGTATATAATTATGAGCAACAAATTCGGAAGCGTGTGACAA 906
 Qy 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrPheSerAlaLeuSerLys 220
 Db 907 ACCTCTTACCAAGCTTGCAATTATAAGTTTGGAAATAGTTGCTGCTCTCTCCAAG 966
 Qy 221 GlnIleGlnIleAlaSerThrAsnGlnPheGluSerProValValLeuIleAsn 240
 Db 967 CAATTCAGATAGCAGTACTAATAATGACAGTTTGAACCTCTGTGTGCTTATAAT 1026
 Qy 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
 Db 1027 GCTCAAAACCAACGAGTACAGTATACCAATGTGTGCTGAGTGTGAACCTCCACATC 1086
 Qy 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaIleMetAspAspValProMetThrGln 280
 Db 1087 GCGTGTGCTGCTAAATAGAAACAATATGGCAGCTATTGATGACGATGTTCTATGGCAG 1146
 Qy 281 SerPheGlyCysGlySerTyrAlaLeu 289
 Db 1147 AGCTTTGGATGTGGAAGTTATGCTATT 1173

RESULT 12

TKU25675 810 bp DNA linear PLN 20-JUL-1995
 LOCUS Trichosanthes kirilowii trichosanthin gene, partial cds.
 DEFINITION U25675
 ACCESSION U25675
 VERSION U25675.1 GI:847831

KEYWORDS Trichosanthes kirilowii (Mongolian snake-gourd)
 SOURCE Trichosanthes kirilowii
 ORGANISM Trichosanthes kirilowii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.

1 (bases 1 to 810)

Bao, Y., Chu, R., Han, J., Zhang, H., Pan, N., Gu, X. and Chen, Z.
 Cloning and sequencing of trichosanthin gene and its expression in

Escherichia coli and tobacco plant

JOURNAL Sci. China. B. 36 (No.6), 669-676 (1993)

REFERENCE 2 (bases 1 to 810)

AUTHORS

TITLE
 JOURNAL

Xu, L.

Direct Submission

Submitted (21-APR-1995) Lin Xu, Institute of Biophysics, Academia
 Sinica, Dept. of Protein Engineering, 15 Datun Road, Chaoyang
 District, Beijing 100101, Peoples Republic of China

FEATURES

source

1. .810

/organism="Trichosanthes kirilowii"

/mol_type="genomic DNA"

/db_xref="taxon:3677"

1. .>810

/codon_start=1

/label=JC1073

/product="trichosanthin"

/protein_id="AAA70096.1"

/db_xref="GI:847832"

/translation="MIRFLVLSLLILTLTTPAVGDSFRLSGARSSSVGVFISNL
 RKALNERRRDIYDIPLESTPLGCGERYALHRYVADETISVAIDRVNRYIMGVAGT
 FYFNEASATEAAKIVFKDMRWKVIPLSGNERLQTAAGKIRENIPILGIPALDSAIT
 TLFYFNANSAAALMWLIQSTSEAAKYFIEQIGKRVKDTFIPSFALISLNSWAL
 SRQIQIASTNNGPESPVLINAEQRVITINVDAGVVTSNIALLENENNA"

ORIGIN

Alignment Scores:

Pred. No.: 1.35e-105 Length: 810
 Score: 1262.00 Matches: 256
 Percent Similarity: 97.04% Conservative: 6
 Best Local Similarity: 94.81% Mismatches: 8
 Query Match: 88.13% Indels: 0
 DB: 8 Gaps: 0

US-09-905-247-1 (1-289) x TKU25675 (1-810)

Qy 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20
 Db 1 ATGATCAGATTCTTAGTCTCTCTTTGCTAATCTCACCTCTTCTTAACTCTGCT 60
 Qy 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
 Db 61 GTGAGGCGCATGTTAGCTTCGGTTATCAGGTGCAAGAGCAGTTCCTATGGAGTTTC 120
 Qy 41 IleserAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
 Db 121 ATTTCAATCTGAGAAAGCTCTTGGAAATGAAGAGACTATATGATATCCCTCTGTA 180
 Qy 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
 Db 181 CGTTCACACTCTCCAGGTTGTGAAAGCTAGCCTACGATTCATCTCAGAAATACCGCAT 240
 Qy 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
 Db 241 GAACACATTTTCAGTGGCCATAGACGTAAGGAACGCTATATTATGGATATCGCGTGC 300
 Qy 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
 Db 301 GATACATCTCTATTTTCAACAGGCTTCTGCAACAGAGCTGCAAAATATGTATTCAA 360
 Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
 Db 361 GACGCTATGCGAAAGTACGCTTCCATATTCTGGCAATTACGAAAGGCTTCAAACTGCT 420
 Qy 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
 Db 421 GCAGGCAAAATCAGGAAATATTCGCTTGGACTCCCAAGCTTTGGACAGTCCCATACC 480
 Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
 Db 481 ACTTGTGTTTACTACAACGCCAATTTCTGCTGCTGCGCACTTATGCTACTCATTCAGTCG 540
 Qy 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200
 Db 541 ACGTCTGAGGCTGCGAGGTATAAATTTATTGAGCAACAAATTCGGAAGCGCGTTGACAA 600

QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrPheLeuSerLys 220
 Db 601 ACCTTCATACCAAGTTTGCAGTTTAAAGTTTGAAAATAGTTGGTCTCTCTCCAGG 660
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
 Db 661 CAAATTACAGATACGAGTACTAATAATAGGACAGTTTGAAGTCTCTGTGCTTATAAAT 720
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
 Db 721 GCTGAACCAACAGGATCAGATCACAATGTTGATGCTGGAGTTGTAACCTCCAAATC 780
 QY 261 AlaLeuLeuLeuAsnArgAsnMetAla 270
 Db 781 GCGTTGCTGCTGAATAGAAAACAATATGGCA 810

RESULT 13
 AY082348
 LOCUS
 DEFINITION Trichosanthin kirilowii trichosanthin precursor (TCS) gene, partial cds.
 ACCESSION AY082348
 VERSION AY082348.1 GI:20977594
 KEYWORDS
 SOURCE Trichosanthin kirilowii (Mongolian snake-gourd)
 ORGANISM
 Trichosanthin kirilowii
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthin.
 1 (bases 1 to 1433)
 Yuan,H., Wang,L., Wang,Y., Liu,T., An,C. and Chen,Z.
 Trichosanthin kirilowii trichosanthin precursor (TCS) gene
 Unpublished
 2 (bases 1 to 1433)
 Wang,L., Yuan,H., Liu,T., An,C. and Chen,Z.
 Direct Submission
 Submitted (05-MAR-2002) National Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Sciences, Peking University, Beijing 100871, China
 REFERENCE
 LOCATION/Qualifiers
 1..1433 /organism="Trichosanthin kirilowii"
 /mol_type="genomic DNA"
 /db_xref="taxon:3677"
 1..>1433
 /gene="TCS"
 1..598
 /gene="TCS"
 597
 /gene="TCS"
 /note="alternative weaker transcriptional start site"
 599..>1433
 /gene="TCS"
 /product="trichosanthin precursor"
 599..623
 /gene="TCS"
 600
 /gene="TCS"
 /note="alternative weaker transcriptional start site"
 624..>1433
 /gene="TCS"
 /note="ribosome inactivating protein"
 /codon_start=1
 /product="trichosanthin precursor"
 /protein_id="AA022782.1"
 /db_xref="GI:20977595"
 translation="MIRFLVLISLTLFLTPAVEGDVSPRLSGATSSSYGVFISNL
 RKALPYEKKYDIPIILRSSLSQRYALILYTNVADETISVAIDVTNVIYMGYAGDI
 SYFFNEASATAAKVFKDARKVTLPSYGNVRLQIAAGKIRENIPGLPALDSALT
 TLFYNNASAAALLVLTQSTAEAAKYFIEQIGKRVDTKTLPSLAIISLNSWSAL
 SKQIQIATNNQFESFVLINAQNRVTIANVAGVVISNIAFLNLRNMA"
 624..692
 /gene="TCS"

mat_peptide 693..1433
 /gene="TCS"
 /product="trichosanthin"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,88e-104 Length: 1433
 Score: 1253.00 Matches: 256
 Percent Similarity: 96.67% Conservative: 5
 Best Local Similarity: 94.81% Mismatches: 9
 Query Match: 87.50% Indels: 0
 DB: Gaps: 0

US-09-905-247-1 (1-289) x AY082348 (1-1433)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrThrProAla 20
 Db 624 ATGATCAGATTCTTAGTCTCTCTTTGCTAAATCTCACCTCTCTTCTTACAGCTCTGCT 683
 QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
 Db 684 GTCGAGGGCGATGTTAGCTTCCTGTTATCTGCTGCAACAGCAGTTCTCTATGAGTTTTC 743
 QY 41 IleSerAsnLeuArgGlyAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
 Db 744 ATTTCAATCTGAGAAAGCTCTTCATATGAAGAAGAACTGTAGCATATCCCTCTGTATA 803
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
 Db 804 CGTTCCTCTCTTTCAGGTTCTCAACGCTACGCATTAACTATCTCACAAATTACGCCGAT 863
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
 Db 864 GAACCATTTTCAGTGGCCATAGACGTAAACGACGTCTATATTATGGGATATCGCGTGT 923
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaIysTyrValPheLys 120
 Db 924 GATATATCTATTTTTCACGAGCTCTTCGACAGAGCTGCAAAATATATATATTCAAA 983
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
 Db 984 GACGCTAAGCGAAAAGTTACGCTTCCATATCTCGAAATTACGAAGGCTTCAAATTGCT 1043
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
 Db 1044 GCAGCAAAATTAAGGAAAAATATTCGCTTGGACTCCCGAGCTTTGGACAGTGCATATACC 1103
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
 Db 1104 ACTTGTGTTTACTACACGCCAAATTCGCTGCTCGGCACTTCTCGTACTCATTAGTCC 1163
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
 Db 1164 ACGGCTGAGGCTGCGAGGTATAAATTTATTGAGCAACAAATTTGGAGCGGTGTTGACAAA 1223
 QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrPheAlaLeuSerLys 220
 Db 1224 ACCTTCTTACCAGCTAGCAATTATAAGTTTGGAAAAATAGTTGCTGCTCTCTCCCAAG 1283
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
 Db 1284 CAATTCAGATAGCGAGTACTAATAATGACAGATTTGAAAGTCTCTGTTGCTTATAAAT 1343
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
 Db 1344 GCTCAAAACCAACAGTACAGTACGCAATCTTAATGCTGGAGTTGTAACTCCCAACATC 1403
 QY 261 AlaLeuLeuLeuAsnArgAsnMetAla 270
 Db 1404 GCGTTTCTGCTAAATAGAAAACAATATGGCA 1433

RESULT 14
 AB039324

```

LOCUS      AB039324                      741 bp      DNA      linear      PLN 19-JUN-2002
DEFINITION Trichosanthes sp. Bac Kan 8-98 TBK gene for Trichobakin, partial
            cds.
ACCESSION  AB039324
VERSION    AB039324.1 GI:7242889
KEYWORDS   Trichosanthes sp. Bac Kan 8-98
SOURCE     Trichosanthes sp. Bac Kan 8-98
ORGANISM   Trichosanthes sp. Bac Kan 8-98
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
REFERENCE  1
            Van Chi, P., Quoc Trung, H., Thuy Ha, N., Chung, W.I. and Binh, L.T.
            Characterization of trichobakin, a type I ribosome-inactivating
            protein from Trichosanthes sp. Bac Kan 8-98
            Biotechnol. Appl. Biochem. 34 (2), 85-92 (2001)
            1172374
REFERENCE  2 (bases 1 to 741)
            Van Chi, P., Thuy Ha, N. and Binh, L.T.
            Direct Submission
            Submitted (28-FEB-2000) Phan Van Chi, Institute of Biotechnology,
            Protein Biochemistry, Hoang Quoc Viet, Hanoi 84-04, Vietnam
            (E-mail: pychi@netnam.vn, Tel: 7561903)
FEATURES   Location/Qualifiers
            1..741
             /organism="Trichosanthes sp. Bac Kan 8-98"
             /mol_type="genomic DNA"
             /db_xref="taxon:118182"
             /clone="pQV-20"
             /sex="female"
             /cell_type="Leaf"
             /dev_stage="early"
             /note="isolated from Province Bac Kan in the North of
             Vietnam august 1998-vector: pET-21d(+)"
            1..741
             /gene="TBK"
             /gene="TBK"
             /standard_name="type 1 ribosome inactivating protein"
             /codon_start=1
             /product="Trichobakin"
             /protein_id="BAA92530.1"
             /db_xref="GI:7242890"
             /translation="DVSFRLSGATSSSYGFISNLKALPYERKLYDIPLLRSLTPGS
             QRYALIHNTYVADETISVALDVTNVVMYRAGDTSYFNEASATEAAKYVFKDAKRK
             VILPSGNYERLQIAAKIRENIPGLPALSADITLFLVYNANSAALMWLQSTSE
             AARYKFIEQIGKRVKDTFLPSLAISLNSWSALSKQIQAISTNNGQFETPVVLINA
             QNQRVTITNDGAVVTSNIALPNNMA"
ORIGIN
Alignment Scores:
Pred. No.:      3,41e-98      Length:      741
Score:          1180.00      Matches:    239
Percent Similarity: 97.98%      Conservative: 3
Best Local Similarity: 96.76%      Mismatches: 5
Query Match:     82.40%      Indels:    0
DB:              Gaps:      0

US-09-905-247-1 (1-289) x AB039324 (1-741)

QY      24 AspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheLeuSerAsn 43
Db      1 GATGTTAGCTTCGTTTATCAGGTGCAACAAGCAGTTCCTATGAGAGTTTTCATTTCCAAAT 60
QY      44 LeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSer 63
Db      61 CTCGAGAAAGCTCTTCATATGAAGGAAACTGTACGATATCCCTCTGTTACGTTCCACT 120
QY      64 LeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThrIle 83
Db      121 CTTCCAGGTTCTCAACGCTAGCATGGATTGATCCATCTCACAAATTCACCGCATGAACCAATT 180
QY      84 SerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAspThrSer 103

```

```

Db      181 TCAGTGGCCATAGACGTAACGAACGCTCTATGTTATGGGATATCGCGCTGGTATACATCC 240
QY      104 TyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLysAspAlaMet 123
Db      241 TATTTTTCACACGAGGCTTCTGCAACAGAACTGCAAAATATGTTATTCAGAGCGCTAAG 300
QY      124 ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143
Db      301 CGAAAGTTACGCTTCCATATTCGCAATTCAGAAAGCTTCAAAATTCGTCGAGGCAAA 360
QY      144 IleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThrThrIlePhe 163
Db      361 ATAAGGGAAATATTCGCTTGGACTCCAGCTTGGACAGTGGCAATTTACCACTTTGTTT 420
QY      164 TyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGlu 183
Db      421 TACTACACGCCAATTCCTGCTGCTGCGCACTTAAGTACTCATTCAGTCAGCGCTGAG 480
QY      184 AlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLysThrPheLeu 203
Db      481 GCTCGAGGTATAAATTTATTGAGCAACAAATTCGGAAGCGTGTGTGACAAACCTTCCTA 540
QY      204 ProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLysGlnIleGln 223
Db      541 CCAAGTCTTGCATTTATAAGTTTGGAAATAGTTGGTCTGCTCTCTCCAGCAAAATTCAG 600
QY      224 IleAlaSerThrAsnAsnGlyInPheGluSerProValValLeuIleAsnAlaGlnAsn 243
Db      601 ATAGGAGTACTAATAATGACAGTTTGAAACTCTCTGTGTGCTTATAAATGCTCAAAAC 660
QY      244 GlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIleAlaLeuLeu 263
Db      661 CAACGAGTCACGATAACCAATGTTGATGCTGGAGTTGTAACCTCCACATCGCGTGTGCTG 720
QY      264 LeuAsnArgAsnAsnMetAla 270
Db      721 CCAATAGAAACAATATGGCA 741

RESULT 15
LOCUS      BYNBD1                      1084 bp      mRNA      linear      PLN 21-NOV-2001
DEFINITION Bryonia dioica ribosome-inactivating protein mRNA, complete cds.
ACCESSION  L42298
VERSION    L42298.1 GI:17028175
KEYWORDS   bryodin 1; ribosome inactivating protein.
SOURCE     Bryonia dioica (red bryony)
ORGANISM   Bryonia dioica
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.
REFERENCE  1 (bases 1 to 1084)
            Gawlak, S.L., Neubauer, M.G. and Siegall, C.B.
            Cloning, expression, and structure-function analysis of bryodin 1,
            a ribosome-inactivating protein from the plant Bryonia dioica
            Unpublished
            2 (bases 1 to 1084)
            Neubauer, M.G.
            Direct Submission
            Submitted (01-DEC-1996) M.G. Neubauer, Molecular Biology,
            Bristol-Myers Squibb Pharm. Res. Inst., 3005 1st Ave., Seattle, WA,
            USA, 98121 neubauer@bms.com
FEATURES   Location/Qualifiers
            1..1084
             /organism="Bryonia dioica"
             /mol_type="mRNA"
             /db_xref="taxon:3652"
             44..916
             /note="precursor
             putative"
             /codon_start=1
             /product="ribosome-inactivating protein"
             /protein_id="AAL34334.1"
CDS

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 14, 2004, 20:00:06 ; Search time 87 Seconds
(without alignments)
1843.458 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MRFLVLSLLILTLFTTPA.....AAMDDVPMTQFGCGSYAL 289

Scoring table:

	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09905247/runat_10092004_103934_4747/app_query.fasta_1.455
-DB=Issued Patents NA -OFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09905247@cgn_1_105@runat_10092004_103934_4747 -NCPU=6 -ICPU=3
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : Issued Patents NA:

```
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6C_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	1413	98.7	886	1	US-07-923-692C-3
2	1413	98.7	886	1	US-08-184-237-3
3	1413	98.7	886	2	US-08-482-920-3
4	1413	98.7	886	3	US-08-484-341-3
5	1413	98.7	886	3	US-08-483-502-3
6	1413	98.7	886	4	US-09-726-851A-3
7	1164.5	81.3	1094	1	US-08-245-754A-1
8	1164.5	81.3	1094	2	US-08-537-731-1
9	1074	75.0	1499	3	US-08-902-486-14
10	1064	74.3	758	3	US-08-902-486-6
11	728	50.8	913	1	US-07-901-707-13
12	728	50.8	913	1	US-07-988-430-13

13	728	50.8	913	1	US-08-425-336-13	Sequence 13, Appl
14	728	50.8	913	1	US-08-488-113B-13	Sequence 13, Appl
15	728	50.8	913	1	US-08-477-484B-13	Sequence 13, Appl
16	728	50.8	913	2	US-08-646-360-13	Sequence 13, Appl
17	728	50.8	913	3	US-08-839-765-13	Sequence 13, Appl
18	728	50.8	913	3	US-09-136-389-13	Sequence 13, Appl
19	728	50.8	913	4	US-09-610-838-13	Sequence 13, Appl
20	728	50.8	913	4	US-09-711-485-13	Sequence 13, Appl
21	728	50.8	913	5	PCT-US92-09487-13	Sequence 13, Appl
22	537.5	37.5	962	1	US-08-324-301-14	Sequence 14, Appl
23	427	29.8	1855	4	US-09-403-752A-31	Sequence 31, Appl
24	427	29.8	1879	4	US-09-147-208-23	Sequence 23, Appl
25	427	29.8	1879	4	US-09-550-117A-23	Sequence 23, Appl
26	425	29.7	1605	2	US-08-356-786-9	Sequence 9, Appl
27	423	29.5	1698	4	US-09-785-921A-1	Sequence 1, Appl
28	423	29.5	1805	4	US-09-785-921A-10	Sequence 10, Appl
29	423	29.5	1855	4	US-09-147-208-33	Sequence 33, Appl
30	423	29.5	1855	4	US-09-147-208-40	Sequence 40, Appl
31	423	29.5	1855	4	US-09-147-208-47	Sequence 47, Appl
32	423	29.5	1855	4	US-09-147-208-54	Sequence 54, Appl
33	423	29.5	1855	4	US-09-550-117A-33	Sequence 3, Appl
34	423	29.5	1855	4	US-09-550-117A-40	Sequence 40, Appl
35	423	29.5	1855	4	US-09-550-117A-47	Sequence 47, Appl
36	423	29.5	1855	4	US-09-550-117A-54	Sequence 54, Appl
37	423	29.5	1855	4	US-09-403-752A-3	Sequence 3, Appl
38	423	29.5	1855	4	US-09-403-752A-5	Sequence 5, Appl
39	423	29.5	1855	4	US-09-403-752A-7	Sequence 7, Appl
40	423	29.5	1855	4	US-09-403-752A-9	Sequence 9, Appl
41	423	29.5	1855	4	US-09-403-752A-13	Sequence 13, Appl
42	423	29.5	1855	4	US-09-403-752A-15	Sequence 15, Appl
43	423	29.5	1855	4	US-09-403-752A-17	Sequence 17, Appl
44	423	29.5	1855	4	US-09-403-752A-19	Sequence 19, Appl
45	423	29.5	1855	4	US-09-403-752A-21	Sequence 21, Appl

ALIGNMENTS

```
RESULT 1
US-07-923-692C-3
; Sequence 3, Application US/07923692C
; Patent No. 5316931
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantham, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann Myers
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CAL
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,692C
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 600,244
; FILING DATE: 22-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 641,617
; FILING DATE: 16-JAN-1991
; PRIOR APPLICATION DATA:
```

→

QY	1	MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrThrProAla	20
DB	8	ATGATCAGATCTTAGTCCTCTCTTTGCTAATCTCACCCCTTCTCTTCAACATCCTGCT	67
QY	21	ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe	40
DB	68	GTGGAGGGCGATGTAGTCTCGGTTTTCAGGTGCCAACAGCAGTTCTCTATGGAGTTTC	127
QY	41	IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysIleuTyrAspIleProLeuLeu	60
DB	128	ATTTCAATCTGAGAAAAGCTCTTCCAAATGAAGGAAACTGTACGATATCCCTCTGTTA	187
QY	61	ArgSerSerLeuProGlySerGlnArgTyrAlaIleuIleHisIleuThrAsnTyrAlaAsp	80
DB	188	CGTTCCTCTCTCCAGGTTCTCAACGTTAGCATTTGATCCATCTCACAATATACGCCGAT	247

RESULT 3
US-08-482-920-3
; Sequence 3, Application US/08482920
; Patent No. 5866785
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantham, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann Myers
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,920
; FILING DATE: 07-JUNE-1995

Wed Sep 15 10:32:02 2004

```

PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 184,237
  FILING DATE: 19-JAN-1994
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 600,244
    FILING DATE: 22-OCT-1990
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 641,617
    FILING DATE: 16-JAN-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 310,881
    FILING DATE: 17-FEB-1989
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 160,766
    FILING DATE: 26-FEB-1988
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 160,771
    FILING DATE: 26-FEB-1988
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 347,637
    FILING DATE: 05-MAY-1989
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 363,138
    FILING DATE: 08-JUN-1989
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 219,279
    FILING DATE: 15-JUL-1988
  ATTORNEY/AGENT INFORMATION:
    NAME: Halluin, Albert P.
    REGISTRATION NUMBER: 28,957
    REFERENCE/DOCKET NUMBER: 8129-112
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 415-854-3660
    TELEFAX: 415-854-3694
  INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 886 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    HYPOTHETICAL: NO
    ANTI-SENSE: NO
    ORIGINAL SOURCE:
      ORGANISM: Chinese cucumber
      IMMEDIATE SOURCE:
        CLONE: alpha-trichosanthin
    FEATURE:
      NAME/KEY: CDS (B) LOCATION: 8..877
      LOCATION: 8..877
  US-08-482-920-3

Alignment Scores:
  Pred. No.: 1,06e-166 Length: 886
  Score: 1413.00 Matches: 285
  Percent Similarity: 99.31% Conservative: 2
  Best Local Similarity: 98.62% Mismatches: 2
  Query Match: 98.67% Indels: 0
  DB: 2 Gaps: 0

US-09-905-247-1 (1-289) x US-08-482-920-3 (1-886)
QY 1 MetileArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20
Db 8 ATGATCAGATTCTTAGTCTCTCTTTGCTAAATTTCTCACCTCTTCTTAACTCTCGCT 67
QY 21 ValGluCluYAspValSerPheArgLeuSerGlyValAlaThrSerSerTyrGlyValPhe 40
Db 68 GTGAGGGCGGATGTAGCTTCCGTTTATCAGGTGCACACAGCAGTCTCTATGGAGTTTC 127
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 128 ATTTCAAATCTGAGAAAAGCTCTTCCAAATGMAAGGAAACTGTACGATATCCCTCTGTTA 187

61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuHisLeuThrAsnTyrAlaAsp 80
188 CGTTCTCTCTCTCCAGGTTCTCAACGCTAGCATTCATCATCTCAAAATACGCCGAT 247
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
248 GAAACCATTTTCAGTGGCCATAGACGTAAACGACGCTATATTAAGGATATCCGCTGGC 307
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
308 GATACATCTCTATTTTTCACAGAGGCTTCTGCAACAGAGCTGCAAAATATGATTCAAA 367
121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
368 GACGCTATCGAAAAAGTTACGCTTCATATCTTGCAATTTACGAAAGCTTCAAACTGCT 427
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
428 GCGGGCAAAATAAGGAAAAATATTCGCTCGGTCGCACTTATGCTACTCATTCAGTCG 487
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
488 ACITTTGTTTACTACAACGCCAATTTCTGCTCGGTCGCACTTATGCTACTCATTCAGTCG 547
QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200
548 ACGTCGAGGCTGCGGTATAAATTTATTGAGCAACAAATTTGGAAAGCGGTTGACAAA 607
QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrPheSerAlaLeuSerLys 220
608 ACCTTCTTACCAGTTTAGCAATTTAAGTTTGGAAATAAGTTGCTGCTCTCTCTCCAG 667
QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
668 CAAATTCAGATAGGAGTACTAATAATGACAGATTGAACTCTCTGTTGCTGCTATAAAT 727
QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
728 GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGGAGTTGTAACTTCAACATC 787
QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspAspValProMetThrGln 280
788 GCGTTGCTGCTGAATCGAAACAAATATGCGACCATGATGACGATGTTCTTCTATGACAG 847
QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
848 ACCTTTGGATGTGGAAGTTATGCTATT 874
Db

RESULT 4
US-08-484-341-3
; Sequence 3, Application US/08484341
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; Dawson, William O.
; Grantham, George L.
; Turpen, Thomas H.
; Turpen, Ann Myers
; Garger, Stephen J.
; Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CAL
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25

```


Db	188	CGTTCTCTCTCTCCAGGTTCTCAACGCTACGCAATTGATCCATCTCACAAATTACGCCGAT	247
QY	81	GlutThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly	100
Db	248	GAACACCAATTCAGTGGCCATAGACGTACGAACGCTCTATATTATGGATATCGCGCTGGC	307
QY	101	AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys	120
Db	308	GATACATCTCTATTTTTTCAACGAGGCTTCTGCAACAGAAGCTGCACAAATATGTATTCAA	367
QY	121	AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla	140
Db	368	GACGCTATCGGAAAGTTACGCTTCCATATTCTGCAATTACGAAAGGCTTCAAACTGCT	427
QY	141	AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr	160
Db	428	CGGGCAAAATAAGGMAAATAATCCGCTTGACTCCAGCTTTGGACAGTGCCATTACC	487
QY	161	ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer	180
Db	488	ACTTTGTTTACTACACGCGCAATTCGCTCGCTCGGCACCTTATGTTACTCATTCAGTCG	547
QY	181	ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys	200
Db	548	ACGTCCTGAGGCTGCGAGGTATTAATTTATTGAGCAACAAATGGGAAGCGGTTGACAAA	607
QY	201	ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys	220
Db	608	ACCTTCTCACCAGTTTACGCAATTTAAGTTTGGAAATAGTTGGTCTGCTCTCTCCAAG	667
QY	221	GlnIleGlnIleAlaSerThrAsnAsnGlyClnPheGluSerProValValLeuIleAsn	240
Db	668	CAAAATTCAGATAGCGAGTACTATAATATGACAGCTTCAAACTCCTGTTGCTTATAAT	727
QY	241	AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle	260
Db	728	GCTCAAAACCAACGAGTCATGATACCAATGTTGATGCTGAGGTGTAAACCTCCAAATC	787
QY	261	AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln	280
Db	788	CGGTTGCTGCTGAATCGAAACAATATGGCAGCCATGGATGACGATGTTCTATGACACAG	847
QY	281	SerPheGlyCysGlySerTyrAlaLeu	289
Db	848	AGCTTTGGATGTGGAAGTTATGCTATT	874
RESULT 5			
US-08-483-502-3			
; Sequence 3, Application US/08483502			
; Patent No. 6284492			
; GENERAL INFORMATION:			
; APPLICANT: Donson, Jon			
; APPLICANT: Dawson, William O.			
; APPLICANT: Grantham, George L.			
; APPLICANT: Turpen, Thomas H.			
; APPLICANT: Turpen, Ann Myers			
; APPLICANT: Garger, Stephen J.			
; APPLICANT: Grill, Laurence K.			
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS			
; NUMBER OF SEQUENCES: 11			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti			
; STREET: 1201 New York Avenue N.W., Suite 1000			
; CITY: Washington			
; STATE: DC			
; COUNTRY: USA			
; ZIP: 20005			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			

RESULT 6
US-09-726-651A-3
; Sequence 3, Application US/09726651A
; Patent No. 6448046
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: DAWSON, William O.
; APPLICANT: GRANTHAM, GEORGE L.
; APPLICANT: TURPEN, Thomas H.
; APPLICANT: TURPEN, Ann M.
; APPLICANT: GARGER, Stephen J.
; APPLICANT: GRILL, Laurence K.
; TITLE OF INVENTION: RECOMBINANT ANIMAL VIRUS NUCLEIC ACIDS
; FILE REFERENCE: 008010023CNUS01
; CURRENT APPLICATION NUMBER: US/09/726, 651A
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 08/483,502
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/184,237
; PRIOR FILING DATE: 1994-01-19
; PRIOR APPLICATION NUMBER: 07/923,692
; PRIOR FILING DATE: 1992-07-31

APPLICATION NUMBER:	US/08/483,502
FILING DATE:	
CLASSIFICATION:	
Prior APPLICATION DATA:	
APPLICATION NUMBER:	07/739,143
FILING DATE:	
APPLICATION NUMBER:	US 600,244
FILING DATE:	22-OCT-1990
Prior APPLICATION DATA:	
APPLICATION NUMBER:	US 641,617
FILING DATE:	16-JAN-1991
Prior APPLICATION DATA:	
APPLICATION NUMBER:	US 310,881
FILING DATE:	17-FEB-1989
Prior APPLICATION DATA:	
APPLICATION NUMBER:	US 160,766
FILING DATE:	26-FEB-1988
Prior APPLICATION DATA:	
APPLICATION NUMBER:	US 160,771
FILING DATE:	26-FEB-1988
Prior APPLICATION DATA:	
APPLICATION NUMBER:	US 347,637
FILING DATE:	05-MAY-1989
Prior APPLICATION DATA:	
APPLICATION NUMBER:	US 363,138
FILING DATE:	08-JUN-1989
Prior APPLICATION DATA:	
APPLICATION NUMBER:	US 219,279
FILING DATE:	15-JUL-1988
ATTORNEY/AGENT INFORMATION:	
NAME:	Ihnen, Jeffrey L.
REGISTRATION NUMBER:	28,957
REFERENCE/DOCKET NUMBER:	18604-090574
TELECOMMUNICATION INFORMATION:	
TELEPHONE:	202-962-4810
TELEFAX:	202-962-8300
INFORMATION FOR SEQ ID NO:	3:
SEQUENCE CHARACTERISTICS:	
LENGTH:	886 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	single
TOPOLOGY:	linear
MOLECULE TYPE:	DNA (genomic)
HYPOTHETICAL:	NO
ANTI-SENSE:	NO
ORIGINAL SOURCE:	
ORGANISM:	Chinese cucumber
IMMEDIATE SOURCE:	
CLONE:	alpha-trichosanthin
FEATURE:	
NAME/KEY:	CDS
LOCATION:	8..877
US-08-483-502-3	
Alignment Scores:	
Pred. No.:	1.06e-166
Score:	1413.00
Percent Similarity:	99.31%
Best Local Similarity:	98.62%
Query Match:	98.67%
DB:	3
US-09-905-247-1 (1-289) x US-08-483-502-3 (1-886)	
	Length: 886
	Matches: 285
	Conservative: 2
	Mismatches: 2
	Indels: 0
	Gaps: 0
QY	1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20
Db	8 ATGATCAGATTCTTAGTCTCTCTTTTGTAATTCTCACCCCTCTTCCTAACAACTCCTGCT 67
QY	21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPhe 40
Db	68 GTGAGGGCGAAGTTAGTCTTCGTTTATCAGTGCAACAGCAGTTCTATGAGATTTC 127
OY	41 IleSerAsnLeuArgGlyAlaLeuProAnGluArgLysLeuTyrAspIleProLeuLeu 60

```

; PRIOR APPLICATION NUMBER: 07/600,244
; PRIOR FILING DATE: 1990-10-22
; PRIOR APPLICATION NUMBER: 07/641,617
; PRIOR FILING DATE: 1991-01-16
; PRIOR APPLICATION NUMBER: 07/737,899
; PRIOR FILING DATE: 1991-07-26
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 886
; TYPE: DNA
; ORGANISM: Chinese Cucumber
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8)...(877)
; OTHER INFORMATION: Clone:alpha-trichosanthin
US-09-726-651A-3

Alignment Scores:
Pred. No.: 1,06e-166 Length: 886
Score: 1413.00 Matches: 285
Percent Similarity: 99.31% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 2
Query Match: 98.67% Indels: 0
DB: 4 Gaps: 0

US-09-905-247-1 (1-289) x US-09-726-651A-3 (1-886)

Qy 1 MetIleAArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrThrProAla 20
Db 8 ATGATCAGATTCTTAGTCCTCTCTTTGGCTAAATCTCACCTCTTCTTAAACAACCTCGCT 67
Qy 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 68 GTGAGGCGGATGTTAGCTTCGGTTATCAGGTGCACCAAGCAGTTCCTATGGAGTTTC 127
Qy 41 IleSerAsnLeuArgGlyAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 128 ATTTCAAATCTGAGAAAAGCTTCCAAATGAAAGGAAACTGTACGATATCCCTCTGTTA 187
Qy 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
Db 188 CGTTCTCTCTTCAGGTTCTCAACGCTACGCAATGGATCCATCTCACAAATACGCCGAT 247
Qy 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db 248 GAAACCATTTTCAGTGGCCATAGACGTACGACGACGCTATATTTATGGGATATCGCGTGGC 307
Qy 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Db 308 GATACATCCCTATTTTTTCAACGAGGCTTCTGCAACAGAAAGCTGCAAAATATGTAATCAA 367
Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 368 GACGCTATGCGAAAAGTTACGCTTCCATATTTCTGGCAATACGAAAGGCTTCAAACTGCT 427
Qy 141 AlsGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 428 GCGGGCAAAATAAAGGAAAATATTCCGCTTGGACTCCCGAGCTTTGGCAGTGCCATTACC 487
Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 488 ACTTTGTTTACTACAAACGCCAAATCTGCTCGCTCGGCACATTATGGTACTCATTCAGTCG 547
Qy 181 ThrSerGluAlaAlaArgTyrLysPheIleGlnGlnIleGlyLysArgValAspLys 200
Db 548 ACGTCTGAGGCTCGGAGGTATAAATTATTAGCAACAATAATGGGAAGCGGTTGACAAA 607
Qy 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
Db 608 ACCTTCTCAACCAAGTTTAGCAATATTAAGTTTGGAAAATAGTTCGTCTCTCTCCAAG 667
Qy 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240

```

QY	21	ValGluGlyAspValSerPheArgLeuSerGlyValAlaThrSerSerSerTyrGlyValPhe	40
Db	104	GITGAGGGCAUGTTAGCTTCGTTTATCAGGTGCTACAACCACTCTATGAGATTTTC	163
QY	41	IleSerAsnLeuArgLysAlaLeuProAsnGluuArgLysLeuTyrAspIleProLeuLeu	60
Db	164	ATTAATAAATCTGAGAGAAGCTCTTCCATACGAAGGAAAGTGTAACAATATACCGCTATT	223
QY	61	ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp	80
Db	224	CGTTCAGTATTTCAGGTTTCCAGCGCTCACATTACTCCATCTCACAAATTACGCGAT	283
QY	81	GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly	100
Db	284	GAACCATCTCAGTGGCAGTAGACGTAAACAAACGCTCTATATTATGGGCTATCTTGGCGGT	343
QY	101	AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaIalysTyrValPheLys	120
Db	344	GATGTGTCCTATTTTTCACGAGGCTTCAGCAACAGAAAGCTGCAGAAATTCGTATTCAAA	403
QY	121	AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluuArgLeuGlnThrAla	140
Db	404	GATGCTTAAGAAAAGTAGCGTCTCCATATTCAGGCAATTCAGAAAGGCTTCAAACTGCT	463
QY	141	AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr	160
Db	464	GCAGGAAAAATTAAGAGAAAATATTCACCTTGAGCTCCAGAGTTTGGACAGTGCCATTACC	523
QY	161	ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer	180
Db	524	ACTTTGTATTACTACCGCAGATTCTCGGCTTCGCACATTCGTGACTCATTCATCAATCC	583
QY	181	ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys	200
Db	584	ACGGCTGAATCTGCAAGGTATAAAATTTATTGAACAACAAATTTGGAAGCGTGTACACAA	643
QY	201	ThrPheLeuProSerLeuAlaIleIleSerLeuGluuAsnSerTrpSerAlaLeuSerLys	220
Db	644	ACTTTTTTACCAGGTTTATAGCAACTATTAGTTTGGAAAAATAATTTGGTCTGCTCTGTC	703
QY	221	GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn	240
Db	704	CAAAATTCAGATAGCCAGTACCAATAATGGACAAATTTGAGAGTCCCTGTGTGCTTATAGAT	763
QY	241	AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle	260
Db	764	GGTAACAACCAACGAGTCTCTATAACCAATGCTAGTGTCTGAGTGTGAACCTCCAAACATA	823
QY	261	AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln	280
Db	824	GCGTTTGCTGCTAAACAAGAAATATATATTCAGCGCATTTGGAGAGGACATTTCTATGACACT	883
QY	281	---SerPheGlyCysGlySerTyrAlaLeu	289
Db	884	ATTCGGCTTTGAACATGCATCTTATGGTATA	913

```

RESULT 8
US-08-597-731-1
; Sequence 1, Application US/08597731
; Patent No. 593247
; GENERAL INFORMATION:
; APPLICANT: Siegal, Clay B.
; TITLE OF INVENTION: Cloning and Expression of a Gene
; TITLE OF INVENTION: Encoding Bryodin 1 From Bryonia dioica
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: 3005 First Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
;

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/597,731
FILING DATE: 07-FEB-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,754
FILING DATE: 06-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Poor, Brian W.
REGISTRATION NUMBER: 32,928
REFERENCE/DOCKET NUMBER: ON0124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 728-4800
TELEFAX: (206) 727-3601
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1094 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Bryonia dioica
US-08-597-731-1

Alignment Scores:
Pred. No.: 1,55e-135 Length: 1094
Score: 1164.50 Matches: 235
Percent Similarity: 90.00% Conservative: 26
Best Local Similarity: 81.03% Mismatches: 28
Query Match: 81.32% Indels: 1
DB: 2 Gaps: 1

US-09-905-247-1 (1-289) x US-08-597-731-1 (1-1094)
QY 1 MetIleArgPheUeuValLeuSerLeuLeuLeuLeuThrLeuPheLeuThrThrProAla 20
Db 44 ATGATCAAAATGTTAGTCCITGGTTGCTAATTCCTCACCATATTCTCAAAATCTCCAAC 103
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 104 GTTCAGGGCGAGTTAGTCTCCGTTTATCAGGTGCTACAAACACATCTATGGAGTTTTC 163
QY 41 IleSerAsnLeuArgLyLeuAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 164 ATTAATAAATCTGAGAGAAAGCTCTTCATACGAAAGAAAGTGTACAAATATACCGCTATT 223
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuHisLeuThrAsnTyrAlaAsp 80
Db 224 CGTTCAGTATTTCAGGTTTCAGGACGCTACACATTAATCTCCTCACAATATTCGCGGAT 283
QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db 284 GAAACCATCTCAGTGGCAGTAGACGTTCATATGAAACAGCTTATATTATGGGGTATCTTGC 343
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Db 344 GATGTGTCCTATTTTTTCAACCGAGGCTTCAGCAACAGAGAGCTGCAAAATTCGTATTCAA 403
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 404 GATGCTTAAGAAAAAGTGCAGCTTCATATTACGGCAATTTACGAAAGGCTTCAAACTGCT 463
QY 141 AlaGlyLyIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 464 GCAGGAAAAAATTAAGAAAAATATTCACCTCGACTCCAGCTTGCAGCAGTGCATATACC 523
QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180

```

```

Db      524  ACTTTGATTACACCGCGAGTTCTGCGCTTCTGCACTTCTTGACTCAATCAATCC 583
QY      181  ThrSerGluAlaAlaArgTyrLysPheIleGlnGlnIleGlyLysArgValAspLys 200
Db      584  ACGGCTGAATCTCAAGATATAATTTATTGTAACAAATTTGAAAGCGGTAGACAAA 643
QY      201  ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
Db      644  ACTTTTACCAAGTTTAGCACTATTAGTTTGAAATAATTTGGTCTGCTGCTCCAAG 703
QY      221  GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db      704  CAATTCAGATACCAAGTACCAATAATGGACAATTTGAGAGTCTGTTGTCTTAGAT 763
QY      241  AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db      764  GGTAACAACCAACGAGTCTCTATAACCAATGCTAGTGTCTGAGTTGTAACCTCCACATA 823
QY      261  AlaLeuLeuLeuAsnArgAsnMetAlaMetAspAspValProMetThrGln 280
Db      824  GCGTTGCTGCTTAACAGAAATAATATTGCGAGCATTTGGAGGACATTTCTATGACATC 883
QY      281  ---SerPheGlyCysGlySerTyrAlaLeu 289
Db      884  ATCGGCTTTGAACATGGACTTTATGGTATA 913

```

RESULT 9

```

US-08-902-486-14
; Sequence 14, Application US/08902486
; Patent No. 6140075
; GENERAL INFORMATION:
; APPLICANT: Russel, David R.
; TITLE OF INVENTION: METHOD FOR PRODUCING ANTIBODIES AND
; TITLE OF INVENTION: PROTEIN TOXINS IN PLANT CELLS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 670513.90261
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1499 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1493
US-08-902-486-14
Alignment Scores:

```

```

Pred. No.: 5,08e-124 Length: 1499
Score: 1074.00 Matches: 217
Percent Similarity: 90.46% Conservative: 20
Best Local Similarity: 82.82% Mismatches: 21
Query Match: 75.00% Indels: 4
DB: 3 Gaps: 1

US-09-905-247-1 (1-289) x US-08-902-486-14 (1-1499)

QY      24  AspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheIleSerAsn 43
Db      6  GATGTGAGCTTTCGTTTATCAGGTGCTACAAACCATCTCTATGAGTTTTCATATAAAAT 65
QY      44  LeuArgGlyAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSer 63
Db      66  CTGAGAGAGCTCTCCATACGAAAGGAAAGTGTACAATATACCGCTATTACGTTCAAGT 125
QY      64  LeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThrIle 83
Db      126  AITTCAGGTTTCAGGAGCTTACACATTTCTCCATCTCACAAATTCGCGGATGAACCATC 185
QY      84  SerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAspThrSer 103
Db      186  TCAGTGGCAGTAGACGTAACAAACGCTATATTATCGGGTATCTTCCCGGTGATGTGTC 245
QY      104  TyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLysAspAlaMet 123
Db      246  TATTTTTCACGAGGCTTTCAGCAACAGAGCTGCAAAATTCGTATTCAAGATGCTAAG 305
QY      124  ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143
Db      306  AAAAAGTACGCTTCCATATTTCAGGCAATTACGAAAGCTTCAAACTCTGCGAGAAA 365
QY      144  IleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe 163
Db      366  ATAAGAGAAAATATTCCACTTGGACTCCCGCTTTGGACAGTGGCATTTACCACTTTGTAT 425
QY      164  TyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGlu 183
Db      426  TACTACACCGCGAGTCTCGCGCTTCTGCACCTTCTGTACTCTTCAATCCACGCGTGA 485
QY      184  AlaAlaArgTyrLysPheIleGlnGlnIleGlyLysArgValAspLysThrPheLeu 203
Db      486  TCTGCAAGGTATAAATTTATTGAACAATAATTTGAAAGCGGTGTAGACAAAACCTTTT 545
QY      204  ProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLysGlnIleGln 223
Db      546  CCAAGTTTAGCAACTATTAGTTTGGAAAATAATTTGGTCTGCTCTGCTGCCAAGCAATTCAG 605
QY      224  IleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsnAlaGlnAsn 243
Db      606  ATAGCCAGTACCAATAATGGAATTTGAGAGTCTCTGTTGTGCTTATAGATGGAACAAC 665
QY      244  GlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIleAlaLeuLeu 263
Db      666  CAACGAGTCTCTATAACCAATGCTAGTGTGAGTTGTAACTCCCAACATAGCGTTGCTG 725
QY      264  LeuAsnArgAsnAsnMetAlaAlaMetAspAsp-----AspValProMetThr 279
Db      726  CTAACAGAAAATAATATTGACGCAATGATGTTACCAAGGCGCATGATGCTGTGTATGACC 785
QY      280  GlnSer 281
Db      786  CAAAAT 791

RESULT 10
US-08-902-486-6
; Sequence 6, Application US/08902486
; Patent No. 6140075
; GENERAL INFORMATION:
; APPLICANT: Russel, David R.
; APPLICANT: Fuller, James T.
; TITLE OF INVENTION: METHOD FOR PRODUCING ANTIBODIES AND

```

Wed Sep 15 10:32:02 2004

TITLE OF INVENTION: PROTEIN TOXINS IN PLANT CELLS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/902,486

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 670513.90261

TELECOMMUNICATION INFORMATION:

TELEPHONE: 608-251-5000

TELEFAX: 608-251-9166

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 758 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 3..748

US-08-902-486-6

Alignment Scores:

Pred. No.: 2,99e-123

Score: 1064.00

Percent Similarity: 93.52%

Best Local Similarity: 85.83%

Query Match: 74.30%

DB: 3

US-09-905-247-1 (1-289) x US-08-902-486-6 (1-758)

QY 24 AspValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsn 43
 Db 6 GATGTGAGCTTTCGTTTATCAGTGTCTACACCATCTCTGAGTTCATTAATAAT 65
 QY 44 LeuArgLysAlaLeuProAsnGluAArgLysLeuTyrAspIleProLeuLeuArgSerSer 63
 Db 66 CTGAGAGAAGCTCTCCATACGAAAGAAAGTGACATATACCGCTATTACGTTCAAGT 125
 QY 64 LeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThrIle 83
 Db 126 ATTTCAGGTTCAGGAGCGCTACACATTACTCTCATCAATTCAGCGATGAACCATC 185
 QY 84 SerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAspThrSer 103
 Db 186 TCAGTGCAGTACAGCTAACACGCTCTATTATGGGTATCTTCCGGGTGATGTGTC 245
 QY 104 TyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLysAspAlaMet 123
 Db 246 TATTTTTCACGAGGCTTCAGCAACAGAGCTGCAAAATTCGTATTCAAGATGCTAAG 305
 QY 124 ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143
 Db 306 AAAAAAGTGAGCTTCCATATTCAGGCATTAACAGAGGCTTCAACTGTCGAGGAAA 365
 QY 144 IleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe 163

Db 366 ATAAGAGAAAATATTCCACTTGAGCTCCAGCTTTGGACAGTCCCATACCACCTTTGTAT 425
 QY 164 TLTyTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGlu 183
 Db 426 TACTACACCGCCAGTTCCTGGGCTTCTGCACTTCTTGTACTCATTCATCCACGGCTGAA 485
 QY 184 AlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLysThrPheLeu 203
 Db 486 TCTGCAAGGTATAAATTTATTGAACACAAATTTGGAAGCGGTGTAGACAAACATTTTITA 545
 QY 204 ProSerLeuAlaIleIleSerLeuGluAsnSerTipSerAlaLeuSerLysGlnIleGln 223
 Db 546 CCAAGTTTACCACTATTAGTTTGGAAAAATAATTTGCTCTGCTCTCCAGCAATTCAG 605
 QY 224 IleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsnAlaGlnAsn 243
 Db 606 ATAGCCAGTACCAATAATGACAAATTTGAGAGTCTCTGTGTGTATATAGATGGTAACAAC 665
 QY 244 GlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIleAlaLeuLeu 263
 Db 666 CAACGAGTCTCTATAACCAATGTAGTCTCGAGTGTGTAGTCTCAACATAGCGTGTCTG 725
 QY 264 LeuAsnArgAsnAsnMetAla 270
 Db 726 CTAACAGAAATAATATGCA 746
 RESULT 11
 US-07-901-707-13
 ; Sequence 13, Application US/07901707
 ; Patent No. 5376546
 ; GENERAL INFORMATION:
 ; APPLICANT: Bernhardt, Susan L.
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Steve F.
 ; APPLICANT: Lane, Julie A.
 ; TITLE OF INVENTION: Materials Comprising and Methods of
 ; USE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins
 ; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; ADDRESSEE: Bicknell
 ; STREET: Two First National Plaza, 20 South Clark
 ; STREET: Street
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/901,707
 ; FILING DATE: 19920619
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/787,567
 ; FILING DATE: 04-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: AG. 5376546and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 27129/30910
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 346-5750
 ; TELEFAX: (312) 984-5750
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 913 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

```

; GENERAL INFORMATION:
; APPLICANT: Bernhard, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Lane, Julie A.
; APPLICANT: Lei, Shau-Ping
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray &
; ADDRESS: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/988,430
; FILING DATE: 19921209
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5416202and, Greta E.
; REGISTRATION NUMBER: 35302
; REFERENCE/DOCKET NUMBER: 31133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-07-988-430-13

Alignment Scores:
Pred. No.: 3.5e-81 Length: 913
Score: 728.00 Matches: 156
Percent Similarity: 72.83% Conservative: 37
Best Local Similarity: 58.87% Mismatches: 70
Query Match: 50.84% Indels: 3
DB: 1 Gaps: 2

US-09-905-247-1 (1-289) x US-07-988-430-13 (1-913)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20
Db 10 ATGGTGAAGTCTTACTCTCTCTTTTAAATATCGCCATCTTCATTGGTGTCTACT 69
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 70 GCCAAGGCGATGTTTACTTCGATTTGTGACTGCCACTGCAAAACCTCACAAATTT 129
QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 130 ATCGAAGATTTCAGGCGACTCTTCCATTTAGCCATAAAGTGTATGATATACCTCTACTG 189
QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80

```

190 TATTCCATATTCCGACTCCAGAGGTTTCATCTCTCGATCTTACAAAGTTATGCATAT 249
81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
250 GAAACCATCTCGGTGCCATAGATGTGACGAACGTTTATGTTGTGGCGTATCGCACCCGC 309
101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
310 GATGTATCTCTACTTTTAAAA--GAATCTCTCTCTGGAAGCTTATAACATCTCTATTCAAA 366
121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
367 ---GGTACGGGAAAAATTACCTGCCATATACCGGTAAATTATGAAAAATCTTCAAACCTGCT 423
141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
424 GCACACAAAAATAAGAGAGAATAATTGATCTTTGGACTCTCTGCTTGTAGTAGTGCCATTACC 483
161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
484 ACATGTTGTTTATACAAATGCCCAATCTGCTCTCTCTGCTTGTGATTTGCTTACTTAATCCAGACG 543
181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
544 ACTGCAGAACTCGAGATTTAGTATATCGACGACACGCTTCTGAAGTATGTTGCCACT 603
201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
604 AACTTTAAGCCAAATCTAGCCATCATAAAGCTTGGAAAAATCAATGCTGCTCTCTCTCCAA- 662
221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
663 CAATATCTTTTGGCGCAGAAATCAAGGAGGAAAAATTTAGAAATCTGTGACCTTATATAAA 722
241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
723 CCTACCGGGGACCGGTTTCAAGTAACCAATGTTGATTCAGATGTTGTAAGAGGTAAATATC 782
261 AlaLeuLeuLeuAsn 265
783 AAACCTCTGCTGAAC 797
RESULT 13
US-08-425-336-13
; Sequence 13, Application US/08425336
; Patent No. 5621083
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,336
; FILING DATE: 18-APR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,691
; FILING DATE: 12-MAY-1993

Db 604 AACCTTAAAGCAATCTAGCCATCATAGCTTGGAAATCAATGGTCTGCTCTCCAA- 662
Qy 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 663 CAATCTTTTGGCGCAGAAATCAAGGAGGAGAAATTTAGAAATCTGTCGACCTTATAAA 722
Qy 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 723 CCTACCGGGAGCGTTCAGTAACCAATGTTGATTTCAGATGTTGAAAAAGGTAATATC 782
Qy 261 AlaLeuLeuLeuAsn 265
Db 783 AAATCTCTGCTGAAC 797

RESULT 14

US-08-488-113B-13
; Sequence 13, Application US/08488113B
; Patent No. 5744580
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,113B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-488-113B-13

Alignment Scores:
Pred. No.: 3.5e-81 Length: 913
Score: 728.00 Matches: 156
Percent Similarity: 72.83% Conservative: 37
Best Local Similarity: 58.87% Mismatches: 70
Query Match: 50.84% Indels: 3
DB: 1 Gaps: 2

US-09-905-247-1 (1-289) x US-08-488-113B-13 (1-913)

Qy 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20
Db 10 ATGGTGAATGCTTACTACTTCTTTTAAATATGCCATCTCAATGGTCTGCTACT 69
Qy 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 70 GCCAAAGGCGATGTTAACTTCGATTTGTCGACTGCCACTGCAAAACCTTACAAATTT 129
Qy 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 130 ATCGAAGATTTCAGGCGGACTCTTCCATTTAGCCATAAAAGTGTATGATATACCTCTACTG 189
Qy 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
Db 190 TATTCACATATTCGACTCCAGACGTTTCATACCTCTCGATCTTACAAAGTTATGCAAT 249
Qy 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db 250 GAAACCATCTCGTGGCCATAGATGTGACGAAAGTTTATGTTGTGGCGTATGCACCCGC 309
Qy 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Db 310 GATGATCTCTACTTTTAAA---GAATCTCTCTCTGAAGCTTATAACATCTCTTCAA 366
Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 367 ---GGTACGCGGAAATTTACATGCTGCTATACCGGTAATTTAGAAAATCTTCAAATGCT 423
Qy 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 424 GCACACAAATAAGAGAGAAATATTGATCTTGACTCCCTGCTTGAGTAGTGCCTATACC 483
Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 484 ACATTTGTTTATTAACAATGCCCAATCTGCTCTCTGCAATGCTGTACTAATCCAGACG 543
Qy 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
Db 544 ACTGAGAGCTGCAAGATTTAGTATATACGAGGACACAGCTTGCTAGTATGTTGCCACT 603
Qy 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
Db 604 AACTTTAAGCCAAATCTAGCCATCATAGCTTGGAAATCAATGCTGCTCTCTCCAA- 662
Qy 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
Db 663 CAATCTTTTGGCGCAGAAATCAAGGAGGAGAAATTTAGAAAATCTGTCGACCTTATAAA 722
Qy 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
Db 723 CCTACCGGGAGCGTTCAGTAACCAATGTTGATTTCAGATGTTGAAAAAGGTAATATC 782
Qy 261 AlaLeuLeuLeuAsn 265
Db 783 AAATCTCTGCTGAAC 797

RESULT 15

US-08-477-484B-13
; Sequence 13, Application US/08477484B
; Patent No. 5756699
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.

Wed Sep 15 10:32:02 2004

APPLICANT: Carroll, Stephen F.
 APPLICANT: Studnka, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 TITLE OF INVENTION: Proteins
 NUMBER OF SEQUENCES: 169
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/477,484B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/425,336
 FILING DATE: 18-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/064,691
 FILING DATE: 12-WAY-1993

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918
 REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A

TELEPHONE: 312/707-8889
 TELEFAX: 312/707-9155
 TELEX: 650 388-1248
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 913 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-477-484B-13

Alignment Scores:
 Pred. No.: 3 5e-81 Length: 913
 Score: 728.00 Matches: 156
 Percent Similarity: 72.83% Conservative: 37
 Best Local Similarity: 58.87% Mismatches: 70
 Query Match: 50.84% Indels: 3
 DB: 1 Gaps: 2

US-09-905-247-1 (1-289) x US-08-477-484B-13 (1-913)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20
 Db 10 ATGGTGAAGGCTACTACTCTCTTTTAAATATCCCAATCTCATTTGGTGTCTTACT 69

QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyGlyValPhe 40
 Db 70 GCCAAGCGGATGTTAACTTCGATTTGTCGACTGCCACTGCAAAACCTACACAAATTT 129

QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyAspIleProLeuLeu 60
 Db 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyAspIleProLeuLeu 60

Db 130 ATCGAAGATTTCAGGCGGACTCTTCATTTAGCCATAAAGTGATGATATACCTCTACTG 189
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuHleSerThrAsnTyrAlaAsp 80
 Db 190 TATTCACATATTTCGACCTCCAGACGTTTCATACCTCTCGATCTTCAAGTTTATGATAT 249
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
 Db 250 GAAACCATCTCGGTGGCCATAGATGACGAAAGCTTTATTTGGGTATTCGACCCCGC 309
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
 Db 310 GATGTATCTCTACTTTTAAAA---GAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 366
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
 Db 367 ---GGTACGCGGAAATTCACATGCCATATACCGGTAATTTATGAAATCTTCAAACTGCT 423
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
 Db 424 GCACACAAAATAAGAGAGAATATTGATCTTGAGACTCCCTGCTTGTAGTAGTGCATTAAC 483
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
 Db 484 ACATTTGTTTATACATGCCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys 200
 Db 544 ACTGAGAAGCTGCAAGATTTAAGTATATCGAGGACACGTTGCTAAGTATGTTGCCACT 603
 QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
 Db 604 AACTTTAAGCCAAATCTAGCCATCATAGCTTGGAAATCAATGGTCTGCTCTCTCCAA- 662
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
 Db 663 CAATCTTTTGGCGCAGAAATTTAGAAATCTCTGTCGACCTTATAAAA 722
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
 Db 723 CCTACCGGGGAAACGGTTTCAAGTAACCAATGTTGATTTCAGATGTTGTAAAGGTAATATC 782
 QY 261 AlaLeuLeuLeuAsn 265
 Db 783 AAATCTCTCTGTAAC 797

Search completed: September 14, 2004, 21:58:14
 Job time : 93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 14, 2004, 17:15:36 ; Search time 385 Seconds

(without alignments)

3188.907 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MIRELVLSLLILTLFLTPA.....AAMDVDPMTQSGGSGVAL 289

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US0905247/runat.10092004.103933.4703/app.query.fasta.1.455
-DB=N Geneseq.29Jan04 -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0905247 @CGN 1.1.708 @runat.10092004.103933.4703 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq.29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1416	98.9	870	AAH41473	Aah41473 Trichosan
2	1413	98.7	886	AAQ65573	Aaq65573 Alpha-tri
3	1413	98.7	886	AAT47094	Aat47094 Chinese c
4	1413	98.7	886	AAT61376	Aat61376 Chinese c
5	1413	98.7	886	AAX29143	Aax29143 Chinese c
6	1413	98.7	886	AAX03385	Aax03385 DNA seque
7	1413	98.7	886	AAI12393	Aai12393 Chinese c
8	1413	98.7	1227	AAQ66343	Aaq66343 Encodes T

9	1413	98.7	1227	2	AAQ31826	Aaq31826 Recombina
10	1413	98.7	1228	2	AAQ26499	Aaq26499 Trichosan
11	1411	98.5	870	6	ABA95171	Aba95171 Native tr
12	1397	97.6	885	2	AAQ37679	Aaq37679 Encodec c
13	1326	92.6	1185	2	AAQ06351	Aaq06351 Alpha-Tri
14	1326	92.6	1185	2	AAQ26505	Aaq26505 Ribosome
15	1214	84.8	800	2	AAQ06346	Aaq06346 Synthetic
16	1214	84.8	800	2	AAQ26500	Aaq26500 Trichosan
17	1211	84.6	751	2	AAQ75305	Aaq75305 Trichosan
18	1164.5	81.3	1094	2	AAT16208	Aat16208 Bryodin-1
19	1074	75.0	1499	4	AAC60429	Aac60429 BDI-628.5
20	1064	74.3	758	4	AAC60425	Aac60425 DNA encod
21	802.5	56.0	1026	2	AAQ32687	Aaq32687 Luffin-f
22	798	55.7	1077	2	AAQ06348	Aaq06348 Alpha-Tri
23	798	55.7	1077	2	AAQ26502	Aaq26502 Ribosome
24	798	55.7	1179	2	AAQ06349	Aaq06349 Alpha-Tri
25	798	55.7	1180	2	AAQ26503	Aaq26503 Ribosome
26	789.5	55.1	1169	2	AAQ06350	Aaq06350 Alpha-Tri
27	789.5	55.1	1169	2	AAQ26504	Aaq26504 Ribosome
28	775	54.1	914	2	AAQ32688	Aaq32688 Luffin-g
29	766.5	53.5	1012	2	AAQ66093	Aaq66093 Luffin co
30	728	50.8	913	2	AAQ42307	Aaq42307 Momordin
31	728	50.8	913	2	AAQ76184	Aaq76184 Type I ri
32	549	38.3	686	2	AAQ26501	Aaq26501 Ribosome
33	546	38.1	686	2	AAQ06347	Aaq06347 Trichosan
34	537.5	37.5	962	2	AAQ89874	Aaq89874 Bryodin-2
35	427	29.8	1855	2	AAQ04222	Aaq04222 PAP-242 i
36	427	29.8	1879	2	AAT97899	Aat97899 Plasmid p
37	426	29.7	1923	1	AAAN1039	Aan91039 DNA seque
38	424.5	29.6	7306	7	ABZ77370	Abz77370 Nucleotid
39	424.5	29.6	8155	7	ABZ77369	Abz77369 Nucleotid
40	423	29.5	1695	1	AAAN60193	Aan60193 Preprotric
41	423	29.5	1698	4	AAI64137	Aai64137 Castor be
42	423	29.5	1805	4	AAI64145	Aai64145 Modified
43	423	29.5	1855	2	AAT97913	Aat97913 Plasmid p
44	423	29.5	1855	2	AAT97919	Aat97919 Plasmid p
45	423	29.5	1855	2	AAT97910	Aat97910 Plasmid p

ALIGNMENTS

RESULT 1

AAH41473

ID AAH41473 standard; DNA; 870 BP.

AC AAH41473;

XX

DT 23-AUG-2001 (first entry)

XX

DE Trichosanthes kirilowii trichosanthin (TCS) nucleotide sequence.

XX

XX Trichosanthes kirilowii; trichosanthin; TCS; mutagenesis; mutation; MTCS;

XX mutant of trichosanthin; Mongolian snake-gourd; bioactivity; selectivity;

XX cancer; virus; HIV; metaphase induced labour; ds.

XX Trichosanthes kirilowii.

XX

PH Key Location/Qualifiers

FT CDS 1..870

FT /*tag= a

FT /product= "trichosanthin (TCS)"

FT /transl_except= (pos:688..690,aa:Gly)

XX

PN CN1283630-A.

XX

PD 14-FEB-2001.

XX

XX 02-AUG-2000; 2000CN-00119553.

XX

XX 02-AUG-2000; 2000CN-00119553.

XX

PA (SHAN-) SHANGHAI INST CYTOBIOLOGY CHINESE ACAD.

XX

PI	Ke Y, Nie H;
XX	WPI; 2001-291745/31.
DR	P-PSDB; AAB99329.
XX	Trichosanthin mutant and its preparing process.
PT	Disclosure; Page 2 (disclosure); 15pp; Chinese.
FX	
FS	The present invention describes a trichosanthin mutant which is prepared through the mutational deformation of the trichosanthin gene and using an expression system. The trichosanthin gene is isolated from Trichosanthes kirilowii (Mongolian snake-gourd). The trichosanthin mutant has several advantages including high bioactivity, high selectivity to target, and strong kill action to cancer cells, virus and HIV. It can also be used for metaphase induced labour. The present sequence represents the nucleotide sequence of wild type trichosanthin which is given in the exemplification of the present invention
XX	
SQ	Sequence 870 BP; 248 A; 193 C; 176 G; 253 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	8.53e-142 Length: 870
Score:	1416.00 Matches: 286
Percent Similarity:	99.31% Conservative: 1
Best Local Similarity:	98.96% Mismatches: 2
Query Match:	98.88% Indels: 0
DB:	Gaps: 0
US-09-905-247-1 (1-289) x AAH41473 (1-870)	
Qy	1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20
Dd	1 ATGATCAGATTCTTAGTCCTCTTTTGCTAATCTCACCCCTTCTTAACAACTCTGTCT 60
Qy	21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Dd	61 GTGGAGGCCGATGTAGCTTCCTCGTTTATCATGGGTGCACAAGCAGTTCCTATGGAGTTTTTC 120
Qy	41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Dd	121 ATTTCAAATCTGAGAAAAGCTCTTCCAATAATGAAGAAGAACTGTACCATATCCCTCTGTTA 180
Qy	61 ArgSerSerLeuProGlySerGlnAtqTyrAlaLeuLeuIleHisLeuThrAsnTyrAlaAsp 80
Dd	181 CGTTCACAGTCTTCACAGTCTCAACGCTACGATTCATCTCAAAATACGCCGAT 240
Qy	81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Dd	241 GAACACATTTTCAGTGGCCATAGACGTAAACAAAGTCTATATTATGGGATATCGCGTGGC 300
Qy	101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Dd	301 GATACATCCTATTTTTCACAGAGGCTTCGCAACAGAGCTCCAAATAATGTATTCAA 360
Qy	121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Dd	361 GACGCTATGCAAAAAGTTACGCTTCATATTCGCAATTCAGAAAGGCTTCAAACTGCT 420
Qy	141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Dd	421 GCAGCAAAATAAGGAAAAATATTCCGCTTGACATCCCTGCTTTGGACAGTGCATTAACC 480
Qy	161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Dd	481 ACTTTGTTTACTACAAAGCAATTTCTCGTCGTCGCACTTATGTGTAATCATTCAGTCG 540
Qy	181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
Dd	541 AGCTCTGAGGCTGCAGGATATAAAATTTATTAGCAACAAATTTGGGAAGCGTGTGACAAA 600
Qy	201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrPheAlaLeuSerLys 220

US-09-905-247-1 (1-289) x AAT61376 (1-886)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrThrProAla 20
 Db 8 ATGATCAGATTCTTAGTCTCTCTTTGCTAATCTCACCCTCTCTCTTAAACAACCTCTGCT 67
 QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrglyValphe 40
 Db 68 GTGGAGGGCGATGTTAGCTTCGCTTTATCAGTGTCAACAAGCAGTTCCTATGAGATTTC 127
 QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyraaplleProLeuLeu 60
 Db 128 ATTTCAAACTGAGAAAGCTCTTCCAAATGAAGGAAACTGTACGATATCCCTCTGTGA 187
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnThrAlaAsp 80
 Db 188 CGTTCCTCTCTCCAGGTTCTCAACGCTACGCTATGATCCATCTCAAAATACGCCGAT 247
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
 Db 248 GAACCAATTCAGTGCCCATAGACGTACGAACTATATATATGCGATATCGCGTGC 307
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
 Db 308 GATACATCTATTTTTCACGAGGCTTCTGCAACAGAGCTGCAAAATATGTATTCAAA 367
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
 Db 368 GACGCTATGCGAAAGTTACGCTTCATATTCGCAATTCGAAAGGCTTCAACTGCT 427
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
 Db 428 GCGGCAAAATAGGGAATAATTCGCTTGACTCCAGCTTTCGACAGTCCCATACC 487
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
 Db 488 ACTTGTGTTTACTACAAGCCCAATTCCTGCTGCGGCACCTTATGCTACTCATTCAGTCG 547
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
 Db 548 ACCTCTGAGGCTGCGAGGTATAAATTTATGAGCAACAATTTGGAGCGGCTTGACAAA 607
 QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTyrSerAlaLeuSerLys 220
 Db 608 ACCCTCTTACCAGTCTAGCAATTAATAGTTTGGAAATAGTTGCTGCTCTCTCCAAG 667
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
 Db 668 CAAATTCAGATAGCGAGTACTAATATGACAGCTTTGAACTCTCTGTTGCTTATAAT 727
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
 Db 728 GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGAGTGTGAACCTCCAAATC 787
 QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
 Db 788 CGGTGCTGCTGAACTGAAACAATATGCGACCATGATGACGATGTTCTTATGACACAG 847
 QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
 Db 848 AGCTTTGGATGGAGCTTATGCTATT 874

RESULT 5

AAAX29143
 ID AAX29143 standard; DNA; 886 BP.

XX AC AAX29143;
 XX AC

DT 20-MAR-2003 (revised)

DT 04-JUN-1999 (first entry)

XX Chinese cucumber alpha-trichosanthin encoding DNA.
 DE

XX KW Recombinant; plant virus; coat protein; systemic infection;
 KW transcription; therapeutic; chinese cucumber; alpha-trichosanthin; ss.
 XX OS Cucumis sp.
 XX PN US5889190-A.
 XX PD 30-MAR-1999.
 XX 07-JUN-1995; 95US-00480432.
 XX 26-FEB-1988; 88US-00160766.
 PR 26-FEB-1988; 88US-00160771.
 PR 15-JUL-1988; 88US-00219279.
 PR 17-FEB-1989; 89US-00310881.
 PR 05-MAY-1989; 89US-00347637.
 PR 08-JUN-1989; 89US-00363138.
 PR 22-OCT-1990; 90US-00600244.
 PR 16-JAN-1991; 91US-00641617.
 PR 26-JUL-1991; 91US-00737899.
 PR 01-AUG-1991; 91US-00739143.
 PR 31-JUL-1992; 92US-00923692.
 PR 19-JAN-1994; 94US-00184237.
 XX (BIOS-) BIOSOURCE TECHNOLOGIES INC.
 XX Garger SJ, Turpen AM, Grill LK, Grantham GL, Donson J, Dawson WO;
 PI Turpen TH;
 DR WPI; 1999-243290/20.
 DR P-PSDB; AAY01374.
 XX Recombinant plant viral nucleic acid derived from a plus sense, single
 PT stranded RNA plant virus - useful for the transcription of products in a
 PT host.
 XX Example 4; Col 43-46; 46pp; English.

XX The invention relates to a recombinant plant viral nucleic acid derived
 CC from a plus sense, single stranded RNA plant virus. The recombinant plant
 CC viral nucleic acid comprises: (a) a first plant viral subgenomic promoter
 CC that is native to the plus sense, single stranded RNA plant virus and
 CC operably joined to a first nucleic acid expression sequence; and (b) a
 CC second plant viral subgenomic promoter that is non-native and is operably
 CC joined to a second nucleic acid expression sequence; where, (i) (a) and
 CC (b) are incapable of recombination with one another; (ii) either the
 CC coat protein coding sequence. The recombinant plant viral nucleic acid
 CC allows the transcription of products in a host, such as therapeutic and
 CC other useful polypeptides or proteins e.g. enzymes, complex biomolecules
 CC and ribosomes. It also gives the option of applying the coding sequence
 CC to the desired organism, tissue, organ or cell, is stable for the foreign
 CC coding sequences and is capable of systemic infection in the plant host.
 CC The transformation and regeneration of target organisms become
 CC unnecessary. (Updated on 20-MAR-2003 to correct PR field.)

XX Sequence 886 BP; 250 A; 199 C; 181 G; 256 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1-83e-141 Length: 886
 Score: 1413.00 Matches: 285

Percent Similarity: 99.31% Conservative: 2

Best Local Similarity: 98.62% Mismatches: 2

Query Match: 98.67% Indels: 0

DB: 2 Gaps: 0

US-09-905-247-1 (1-289) x AAX29143 (1-886)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrThrProAla 20
 Db 8 ATGATCAGATTCTTAGTCTCTCTTTGCTAATCTCACCCTCTCTCTTAAACAACCTCTGCT 67

21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
68 GTGGAGGGCGATGTTAGCTTCGGTTTATCAGGTGCAACAGCAGTTCCTATGGAGTTTTC 127
41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
128 ATTTCAAATCTGAGAAAGCTTCCAAATGAAAGAAAGAACTGTACGATATCCCTCTGTGA 187
61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
188 CGTTCCTCTTCAGGTTCTCAACGCTACGCTATGATCATCTCAAAATTACGCCGAT 247
81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
248 GAAACCAATTCAGTGGCCATAGACGTAAACAGCTCTATATTATGGATATCGCGCTGGC 307
101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaIleTyrValPheLys 120
308 GATACATCCTATTTTTTCAACGAGGCTTCTGCAACAGAGCTGCAAAATATGTATTCAA 367
121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
368 GACGCTATGCGAAGTTACGCTTCCATATTCTGCAATACGAAAGCTTCAAACGTCT 427
141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
428 GCGGGCAAAATAAGGGAATAATTTCCGCTTGGACTCCAGCTTTGGACAGTGGCCATTACC 487
161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
488 ACTTGTGTTTACTACACGCCAATTTCTGCTGGTGGCTGACCTATGTTACTCATTCAGTCG 547
181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
548 ACGTCTGAGGCTGGAGGTATATAATTTATTCAGCAACAAATTTGGGAAGCGGCTTGACAAA 607
201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
608 ACCTCTTACCAAGTTTACCAATTAAGTTTGGAAATAGTTGTGCTGCTGCTCTTCCAG 667
221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
668 CAAATTCAGATAGCGGAGTACTAATATGACAGATTTGAACCTCTGTGCTTATAAT 727
241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
728 GCTCAAAACCAACGAGTCATGATACCAATGTTGATGCTGGAGTTGTACCTCCAAATC 787
261 AlaLeuLeuLeuAsnArgAsnMetAlaAlaMetAspAspValProMetThrGln 280
788 GCGTTGCTGCTGAATCGAAACAAATATGGCAGCCATGGATGACGATGTTCTATGACACAG 847
281 SerPheGlyCysGlySerTyrAlaLeu 289
848 ACCTTTGGATGCGAAAGTTATGCTATT 874

RESULT 6
AA03385
ID AAX03385 standard; DNA; 886 BP.
XX
AC AAX03385;
XX
27-AUG-2003 (revised)
DT 01-APR-1999 (first entry)
XX
DE DNA sequence encoding chinese cucumber alpha-trichosanthin.
XX
KW Chinese cucumber; alpha-trichosanthin; plant virus;
KW RNA plant virus promoter; systemic infection; foreign gene expression;
KW AIDS therapeutic drug; ss.
XX
OS Momordica cochinchinensis.

Location/Qualifiers
8. .877
/*tag= a
/product= "alpha-trichosanthin"

US5866785-A.
02-FEB-1999.
07-JUN-1995;
26-FEB-1988;
26-FEB-1988;
15-JUL-1988;
17-FEB-1989;
05-MAY-1989;
08-JUN-1989;
22-OCT-1990;
16-JAN-1991;
26-JUL-1991;
01-AUG-1991;
31-JUL-1992;
19-JAN-1994;
95US-00482920.
88US-00160766.
88US-00160771.
88US-00219279.
89US-00310881.
89US-00347637.
89US-00363138.
90US-00600244.
91US-00641617.
91US-00737899.
91US-00739143.
92US-00923692.
94US-00184237.

(BIOS-) BIOSOURCE TECHNOLOGIES INC.
Garger SJ, Turpen AM, Grill LK, Grantham GL, Donson J, Dawson WO;
Turpen TH;
WPI; 1999-142035/12.
P-ESDB; AAW84292.

Recombinant plant viral vector - that is capable of systemic infection in
host plant and stable production of heterologous DNA useful for producing
therapeutic proteins for treating e.g. AIDS.

Example 4; Col 43-46; 45pp; English.

The present sequence encodes chinese cucumber alpha-trichosanthin. The
nucleic acid sequence can be expressed in the plant viral constructs of
the invention. The specification describes a recombinant plant viral
nucleic acid derived from a positive (+)-sense RNA plant virus comprising
a native (+)-sense RNA plant virus promoter that is linked to an
expression sequence and a heterologous (+)-sense RNA plant virus promoter
that is linked to an expression sequence. The promoters are incapable of
recombination with each other, and one of the expression sequences
encodes a plant viral coat protein while the other is optionally a
heterologous coding sequence. The plant viral nucleic acid is capable of
systemic infection in a host plant. The viral construct is useful for the
introduction and expression of non-viral foreign genes in plants and the
production of e.g. potential AIDS therapeutic drugs. (Updated on 27-AUG-
2003 to correct OS field.)

Sequence 886 BP; 250 A; 199 C; 181 G; 256 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.83e-141 Length: 886
Score: 1413.00 Matches: 285
Percent Similarity: 99.31% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 2
Query Match: 98.67% Indels: 0
DB: 2 Gaps: 0

US-09-905-247-1 (1-289) x AAX03385 (1-886)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20
Db 8 ATGATCAGATTCCTAGTCCTCTCTTGTCTTAATCTCACCCCTCTTCTTACAACTCTGCT 67
QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 68 GTGGAGGGCGATGTTAGCTTCCGTTTATCAGGTGCAACAGCAGTTCCTATGGAGTTTTC 127

QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
 Db 128 ATTTCAATCTGAGAAAGCTCTTCCAAATGAAAGGAAACTGTACGATATCCCTCTGTTA 187
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80
 Db 188 CGTTCCTCTCTCCAGGTTCTCAACGCTGCGCATTCATCCATCAAAATTAACGCGAT 247
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
 Db 248 GAAACCATTTCAAGTCCATAGACGTAACGAAAGCTCTATATATATGGATATCGCGTGGC 307
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
 Db 308 GATACATCTCTATTTTTCACGAGGCTTCTGCAACAGAGCTGCAAAATATGATTCAA 367
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
 Db 368 GACGCTATGCGAAAGTTACGCTTCCATATTCGGCAATTAACGAAAGGCTTCAAAC 427
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
 Db 428 GCGGCAAAATTAAGGAAATATTCGCTTGGACTCCAGCTTTGGACAGTGCATTAAC 487
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
 Db 488 ACTTGTGTTTACTACAGCCCAATTTCTGCTGGCTGCGCACTTATGCTACTCATTCAGTCG 547
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
 Db 548 AGCTCTGAGGCTGCGAGGTATTAATTTATTTAGCAACAAATTTGGGAAGCGGTTGACAAA 607
 QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrSerAlaLeuSerLys 220
 Db 608 ACCTTCTTACCAAGTTTACCAATTAAGTTTGGAAATAGTTGCTGCTCTCTCCAG 667
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
 Db 668 CAATTCAGATAGCGAGTACATAATATGACAGATTTGAAACTCTCTGTTGCTTATAAT 727
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
 Db 728 GCTCAAAACCAACGATCATGATACCAATGTTGATGCTGGAGTTGTAACTCCCTCAACATC 787
 QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
 Db 788 GCGTGTGCTGTAATCGAAACAAATATGCGCCATGATGATGATGATGATGATGATGATG 847
 QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
 Db 848 AGCTTGGATGGAAGTTATGCTATT 874

RESULT 7

AAAL2393

ID AAA12393 standard; DNA; 886 BP.

XX AC AAA12393;

XX AC AAA12393;

DT 06-AUG-2003 (revised)

DT 24-AUG-2000 (first entry)

XX XX

XX XX

XX Chinese cucumber alpha-trichosanthin DNA.

KW Animal RNA virus; viral coat protein; plant; male sterility; interleukin;

KW EPO; erythropoietin; CSF; colony stimulating factor; Factor VIII; hGH;

KW human growth hormone; melanin; insulin; vaccine;

KW stereo specific catalysis; alpha-trichosanthin; ss.

XX Trichosanthin kirilowii.

OS OS

XX XX

FH Key

FT Location/Qualifiers

FT CDS

FT 8..877

FT /*tag= a

/product= "Alpha-trichosanthin"

US6054566-A.

25-APR-2000.

07-JUN-1995;

95US-00484341.

26-FEB-1988;

88US-00160766.

26-FEB-1988;

88US-00160771.

15-JUL-1988;

89US-00219279.

17-FEB-1989;

89US-00310881.

05-MAY-1989;

89US-00347637.

08-JUN-1989;

89US-00363138.

22-OCT-1990;

90US-00600244.

16-JAN-1991;

91US-00641617.

26-JUL-1991;

91US-00737899.

01-AUG-1991;

91US-00739143.

31-JUL-1992;

92US-00923692.

(BIOS-) BIOSOURCE TECHNOLOGIES INC.

Garger SJ, Grill LK, Turpen TH, Grantham GL, Dawson WO;

Turpen AM, Donson J;

WPI: 2000-338510/29.

P-PSDB; AAY87791.

Recombinant non-retroviral nucleic acid for producing proteins such as

interleukins, melanin and vaccines, comprises subgenomic promoters linked

to sequences coding for viral coat protein and heterologous proteins.

Example 4; Col 57-60; 51pp; English.

This invention describes a novel recombinant viral nucleic acid (I) from

a non-retroviral (+) sense, single stranded animal RNA virus comprising a

nucleic acid sequence coding for a viral coat protein regulated by a

native subgenomic promoter and other two heterologous nucleic acid

sequences regulated by two other subgenomic promoters. (I) is useful for

expressing foreign genes e.g. genes inducing male sterility in plants.

(I) is also useful for producing proteins such as interleukins, EPO

(erythropoietin), CSF (colony stimulating factor), Factor VIII, hGH

(human growth hormone), melanin, insulin, vaccines etc., and enzymes that

are useful for stereo specific catalysis of organic compounds. (I) is

stable and transcribed systemically. The dual subgenomic promoter system

reduces the frequency of recombination thus reducing regeneration of the

wild type virus. This sequence encodes a chinese cucumber alpha-

trichosanthin protein which is described in the method of the invention.

(Updated on 06-AUG-2003 to correct OS field.)

Sequence 886 BP; 250 A; 199 C; 181 G; 256 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,83e-141

Score: 1413.00

Percent Similarity: 99.31%

Best Local Similarity: 98.62%

Query Match: 98.67%

DB: 3

US-09-905-247-1 (1-289) x AAA12393 (1-886)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrProAla 20

Db 8 ATGATCAGATTTCTAGTCTCTCTTGTCTAATTCACCTCTCTCTCTCTCTCTCTCTCT 67

QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40

Db 68 GTGGAGGGCGATGTTAGCTTCGCTTTATCAGGTGCAACAGCAGTTCCTATGAGTTTC 127

QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60

Db 128 ATTTCAATCTGAGAAAGCTCTTCCAAATGAAGGAAACTGTACGATATCCCTCTGTTA 187

QY	61	ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp	80
DB	188	CGTTCCTCTCTCTCCAGGTTCCTCAACGCTAGCATTGATCCATCTCAACAATTACGCCGAT	247
QY	81	GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly	100
DB	248	GAACCATTTTCAGTGCCATAGACGTAAACAAAGCTCTATATTATGGGATATCGCGTGGC	307
QY	101	AspThrSerTyrPhePheHisGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys	120
DB	308	GATACATCTCTATTTTCAACAGAGGCTTCCTGCACACAGAGCTGCAAAATATGATTCAAA	367
QY	121	AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla	140
DB	368	GAGCCTATGCAAAAGTTACGCTTCCATATCTTGGCAATTACGAAAGGCTTCAAACTGCT	427
QY	141	AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaIleAspSerAlaIleThr	160
DB	428	GCGGGCAAAATAAGGGAAAAATATTCGCTTGGACTCCAGCTTTGGACAGTGGCCATTACC	487
QY	161	ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer	180
DB	488	ACTTTGTTTACTACACGCCAAATCTGCTGCGTCGCACATTATGGTACTCATTCAGTCG	547
QY	181	ThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLysArgValAspLys	200
DB	548	ACGTCCTGAGGCTGCGAGGTATAAATTTATTGACCAACAAATTTGGAAAGCGGTTGCACAA	607
QY	201	ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTyrPheSerAlaLeuSerLys	220
DB	608	ACCTTCCTACCAAGTTTAGCAATATATAAGTTTGGAAATAGTTGGTCTGCTCTCTCCAAG	667
QY	221	GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn	240
DB	668	CAAAATTCAGATAGCGAGTACTAAATAATGACAGAGTTTGAAACTCTCTGTTGTGCTTATAAT	727
QY	241	AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle	260
DB	728	GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGGAGTTGTAACTTCCAACTC	787
QY	261	AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln	280
DB	788	GC GTTGTCTGCTGAATCGAAACAAATATGGCAGCCATGGATGACGATGTTTCTTATGACAG	847
QY	281	SerPheGlyCysGlySerTyrAlaLeu	289
DB	848	AGCTTTGGATGTGGAAGTTATGCTATT	874

RESULT 8

AAQ06343	standard; DNA; 1227 BP.
XX	
XX	
XX	AAQ06343;
XX	
DT	06-FEB-1991 (first entry)
XX	
XX	Encodes Trichosanthin from Trichosanthus kirilowii.
DE	
XX	trichosanthin (TCS); Human Immunodeficiency Virus; inhibition; ss.
XX	
XX	Trichosanthes kirilowii.
XX	
XX	Key Location/Qualifiers
PH	340..1209
FT	CDS
FT	/*tag= a
FT	409..1209
FT	mat_peptide
FT	/*tag= b
FT	/product= "TCS"
FT	
XX	
XX	WO9012097-A.
PN	
XX	
PD	18-OCT-1990.

XX 04-APR-1989; 89US-00333184.
XX
XX 04-APR-1989; 89US-00333184.
XX
XX (GENE-) GENELABS INC.
XX
XX Piatek M, Chow T, Fry K;
PI
XX WPI; 1990-334847/44.
XX
XX P-PSDB; AAR07514.
XX
XX Recombinant tri:choanthin protein - with selective inhibitory effect on
PT viral expression in HIV infected T-cells or monocyte-macrophase.
PT
XX Claim 2; Page 59; 102pp; English.
XX
XX Genomic DNA was isolated from T.kirillowii leaves from Korea and a library
CC was constructed. Clone p021D was identified as likely to contain a TCS-
CC encoding sequence in its 4th insert. The deduced amino acid sequence is
CC identical to that of TCS purified from Cantonese T.kirillowii roots,
CC except for 2 conservative substitu-tions. The isolated coding sequence
CC can be used to produce TCS in pure form for use as a therapeutic agent.
CC See also AAQ06344-Q06351
XX
XX Sequence 1227 BP; 402 A; 250 C; 223 G; 352 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 2,85e-141 Length: 1227
Score: 1413.00 Matches: 285
Percent Similarity: 99.31% Conservative: 2
Best Local Similarity: 98.62% Mismatches: 2
Query Match: 98.67% Indels: 0
DB: 2 Gaps: 0

US-09-905-247-1 (1-289) x AAQ06343 (1-1227)

Qy 1 MetIleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20
Db 340 ATGATCAGATCTTAGGCTCTCTTTGCTTAATCTCACCTCTCTCTTACCACTCTGCT 399
Qy 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
Db 400 GTGAGGGCGAGTGTAGCTTCGGTTTATCAGGTGCAACAGCAGTTCCTATGAGGTTTC 459
Qy 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
Db 460 ATTTCAAATCTGAGAAAAGCTCTTCCAAATGAAAGGAACATGACGATATCCCTCTGTTA 519
Qy 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuIleHisLeuThrAsnTyrAlaAsp 80
Db 520 CGTTTCCTCTCTTCAGGTTCCTCAACGCTAGCATTGATTCATCTCACAANTAGCCGAT 579
Qy 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
Db 590 GAACACATTTCAGTGGCCATAGACGTAAACGACCTCTATATTATGGGATATCGCGCTGC 639
Qy 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Db 640 GATACATCCTATTTTTTCAACGAGGCTTCTGCAACAGAGCTGCAAAATATGTATTCAAA 699
Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 700 GACGCTATGCGAAAGTTTACGCTTCCATATTTCTGGCAATTACGAAAGCCTCAAACTGCT 759
Qy 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db 760 GCGGCGCAAAATAAGGAAAAATATTCCGCTTGACCTCCAGCTTTGGACAGTGCATTACC 819
Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
Db 820 ACTTTGTGTTTACTACAGCCCAATTTCTGCTCGTCGGCACTTATGCTACTCATTCAGTCG 879

QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
 Db 880 ACCTCTGAGCTCGAGGTATAAATTTATTGAGCAACAATTCGGAAGCGGTGACAAA 939
 QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
 Db 940 ACCTTCTACCAAGTTTAGCAATATAAGTTTGGAAATAGTTGGTCTGCTCTCTCCAAG 999
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
 Db 1000 CAAATTCAGATAGCGAGTACTAATAATGACAGCTTGAACACTCTGTGTGCTTATAAAT 1059
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
 Db 1060 GCTCAAAACCAACGAGTCATGATAACCAATGTTGATGCTGGAGTTGTAACCTCCAACATC 1119
 QY 261 AlaLeuLeuLeuAsnArgAsnMetAlaAlaMetAspAspValProMetThrGln 280
 Db 1120 GCGTGTCTGCTGATGCAACCAATATGGCAGCCATGGATGACGATGTTCTATGACACAG 1179
 QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
 Db 1180 AGCTTTGGATGTGGAAGTTATGCTATT 1206
 RESULT 9
 ID AAQ31826 standard; DNA; 1227 BP.
 AC AAQ31826;
 DT 25-MAR-2003 (revised)
 DT 16-APR-1993 (first entry)
 XX Recombinant TCS sequence (pQ21D).
 KW TCS; alpha-trichosanthin; Radix-trichosanthin; primer; inhibition;
 KW viral expression; HIV; T-cell; macrophage; ss.
 OS Trichosanthin kirilowii.
 FH Key Location/Qualifiers
 FT CDS 340..1209
 FT /tag= a
 FT /label= TCS
 FT 342..410
 FT /tag= b
 FT /note= "putative N-terminal extension of the mature TCS"
 FT 411..1151
 FT /tag= c
 FT /note= "Claim 3; page 36"
 FT 1152..1208
 FT /tag= d
 FT /note= "putative C-terminal extension of the mature TCS"
 FT US5166056-A.
 PN XX
 PN 24-NOV-1992.
 XX
 XX 09-DEC-1991; 91US-00804293.
 PF XX
 PR 04-APR-1989; 89US-00333184.
 PR 07-SEP-1989; 89US-00404326.
 XX
 PA (GENE-) GENELABS INC.
 XX
 XX Piatak M, Chow TP;
 PI XX
 DR WPI; 1992-414954/50.
 DR P-PSDB; AAR29272.
 XX
 XX Recombinant Trichosanthin protein prodn. in E. coli - for use in the
 PT selective inhibition of viral expression in HIV infected cells.
 XX

PS Disclosure; Fig 4; 37pp; English.
 XX
 CC The TCS coding sequence was amplified using the primers of AAQ31828-30.
 CC The amplified prod. has the sequence of AAQ31827, which was used as a
 CC probe. One isolate, pQ21D, comprises the sequence of AAQ31826. The
 CC recombinant TCS sequence may be used in the recombinant prodn. of TCS.
 CC TCS can be used for the selective inhibition of viral expression in HIV-
 CC infected human T-cells or macrophages. (Updated on 25-MAR-2003 to correct
 CC PF field.)
 XX
 SQ Sequence 1227 BP; 402 A; 250 C; 223 G; 352 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,85e-141 Length: 1227
 Score: 1413.00 Matches: 285
 Percent Similarity: 99.31% Conservative: 2
 Best Local Similarity: 98.62% Mismatches: 2
 Query Match: 98.67% Indels: 0
 DB: 2 Gaps:
 US-09-905-247-1 (1-289) x AAQ31826 (1-1227)
 QY 1 MetIleArgPheLeuValLeuSerLeuLeuIleLeuThrLeuPheLeuThrThrProAla 20
 Db 340 ATGATCAGATCTTAGTCTCTCTTTGCTAATTCACCTCTTCTTACCAACTCTCTGCT 399
 QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
 Db 400 GTGAGGGCGATGTTAGCTTCGTTTATCAGGTGCAACAGCAGTTCCTATGAGGTTTC 459
 QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
 Db 460 ATTTCAAATCTGAGAAAAGCTCTTCCAATGAAGGAAACTGTACGATATCTCTGTTA 519
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrIleAsp 80
 Db 520 CGTTCCTCTCTCCAGGTTCTCAACGCTACCATTCATTCATCTCACAATTTACCCCGAT 579
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
 Db 580 GAAACCATTTTCAGTGGCCATAGACGTACGAAACGCTATATATATGGGATATCGCGCTGC 639
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
 Db 640 GATCATCTCTATTTTTCACGAGCTTCTCAACAGAGCTGCAAAATATATGTTATCAAA 699
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGluThrAla 140
 Db 700 GACGCTATGGGAAAAGTTACGCTTCATATTTCTGGCAATTCAGAAAGCTTCAACCTGCT 759
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
 Db 760 GCGGGCAAAATAGGAAATATATTCGCTTGGACTCCAGCTTTGGACAGTCCCATACC 819
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
 Db 820 ACITTTGTTTACTACAACGCCAATTCGTGCGCTGGCACTTATGTTACTCATTCAGTCG 879
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
 Db 880 ACGTCTGAGGCTCGAGGTATATAATTTATTGAGCAACAATTTGGAAAGCGGTGACAAA 939
 QY 201 ThrPheLeuProSerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
 Db 940 ACCTTCTACCAAGTTTAGCAATATAAGTTTGGAAATAGTTGGTCTGCTCTCCAAG 999
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
 Db 1000 CAAATTCAGATAGCGAGTACTAATAATGACAGCTTGAACACTCTGTGTGCTTATAAAT 1059
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
 Db 1060 GCTCAAAACCAACGAGTCTATGATAACCAATGTTGATGCTGGAGTTGTAACCTCCAACATC 1119

ABA95171 standard; cDNA; 870 BP.

ABA95171;

20-MAY-2002 (first entry)

Native trichosanthin (TCS) encoding cDNA.

Trichosanthin; TCS; mutant; MTCS; therapeutic; cytostatic; antitumour;
anti-human immunodeficiency virus; virucide; immunostimulant; gene;
ectopic pregnancy; ss.

Trichosanthes kirilowii.

Key Location/Qualifiers

CDS 1..870

/tag= a

/transl_except= (pos: 16..18, aa: Val)

/transl_except= (pos: 523..525, aa: Met)

/transl_except= (pos: 865..867, aa: Leu)

/product= "TCS"

sig_peptide 1..69

/tag= b

mat_peptide 70..867

/tag= c

WO200212537-A2.

14-FEB-2002.

18-JUL-2001; 2001WO-CN001178.

02-AUG-2000; 2000CN-00119553.

18-JAN-2001; 2001CN-00103102.

(BEIJ-) BEIJING STM BIOTECH LTD.
(KEYY/) KE Y.
(NIEH/) NIE H.

Ke Y, Nie H;

WPI; 2002-227165/28.

P-PSDB; ABB07660.

Mutant trichosanthin protein of low antigenicity useful for treating
tumor e.g. leukemia, comprises a sequence of native trichosanthin with
modification of at least one amino acid residue in three specific
regions.

Example 1; Fig 1; 42pp; English.

The invention relates to a mutant trichosanthin (MTCS) protein of low
antigenicity comprising a sequence of native TCS with the modification of
at least one amino acid residue in three regions which is 174-180, 203-
227 and 230-244 and substantially retaining the biological activities of
the native TCS. The MTCS protein, its fragment or derivative is useful as
an therapeutic agent; or for preparing a medicament for treating viral
disease e.g. acquired immunodeficiency syndrome (AIDS), tumour e.g.
leukemia, for inducing abortion and/or treating ectopic pregnancy. The
present sequence represents the native TCS encoding cDNA

SQ Sequence 870 BP; 249 A; 192 C; 176 G; 253 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2..92e-141 Length: 870

Score: 1411.00 Matches: 285

Percent Similarity: 99.31% Conservative: 2

Best Local Similarity: 98.62% Mismatches: 2

Query Match: 0 Indels: 0

DB: 6 Gaps: 0

US-09-905-247-1 (1-289) x ABA95171 (1-870)

ID ABA95171 standard; cDNA; 870 BP. 20
AC ABA95171; 20
XX 20-MAY-2002 (first entry) 60
DT Native trichosanthin (TCS) encoding cDNA. 40
XX Trichosanthin; TCS; mutant; MTCS; therapeutic; cytostatic; antitumour; 40
DE anti-human immunodeficiency virus; virucide; immunostimulant; gene; 40
KW ectopic pregnancy; ss. 120
XX Trichosanthes kirilowii. 60
OS Key Location/Qualifiers 180
FH CDS 1..870 80
FT /tag= a 240
FT /transl_except= (pos: 16..18, aa: Val) 100
FT /transl_except= (pos: 523..525, aa: Met) 300
FT /transl_except= (pos: 865..867, aa: Leu) 360
FT sig_peptide 1..69 120
FT mat_peptide 70..867 120
FT /tag= b 360
FT /tag= c 140
PN WO200212537-A2. 140
XX 14-FEB-2002. 140
PD 18-JUL-2001; 2001WO-CN001178. 160
PF 02-AUG-2000; 2000CN-00119553. 160
PR 18-JAN-2001; 2001CN-00103102. 180
XX (BEIJ-) BEIJING STM BIOTECH LTD. 180
FA (KEYY/) KE Y. 180
PA (NIEH/) NIE H. 180
XX Ke Y, Nie H; 200
PI WPI; 2002-227165/28. 200
XX P-PSDB; ABB07660. 200
DR Mutant trichosanthin protein of low antigenicity useful for treating 220
PT tumor e.g. leukemia, comprises a sequence of native trichosanthin with 220
PT modification of at least one amino acid residue in three specific 220
XX regions. 220
XX Example 1; Fig 1; 42pp; English. 220
XX The invention relates to a mutant trichosanthin (MTCS) protein of low 220
CC antigenicity comprising a sequence of native TCS with the modification of 220
CC at least one amino acid residue in three regions which is 174-180, 203- 220
CC 227 and 230-244 and substantially retaining the biological activities of 220
CC the native TCS. The MTCS protein, its fragment or derivative is useful as 220
CC an therapeutic agent; or for preparing a medicament for treating viral 220
CC disease e.g. acquired immunodeficiency syndrome (AIDS), tumour e.g. 220
CC leukemia, for inducing abortion and/or treating ectopic pregnancy. The 220
CC present sequence represents the native TCS encoding cDNA 220
XX SQ Sequence 870 BP; 249 A; 192 C; 176 G; 253 T; 0 U; 0 Other; 220
XX Alignment Scores: 220
XX Pred. No.: 2..92e-141 Length: 870 220
XX Score: 1411.00 Matches: 285 220
XX Percent Similarity: 99.31% Conservative: 2 220
XX Best Local Similarity: 98.62% Mismatches: 2 220
XX Query Match: 0 Indels: 0 220
XX DB: 6 Gaps: 0 220
XX US-09-905-247-1 (1-289) x ABA95171 (1-870) 220

PI	Piatek M, Chow T, Fry K;
XX	WPI; 1990-334847/44.
DR	P-PSDB; AAR07523.
XX	
PT	Recombinant tri:chosanthin protein - with selective inhibitory effect on
PT	viral expression in HIV infected T-cells or monocyte-macrophase.
XX	
PS	Example; Fig 20; 102pp; English.
XX	
CC	pQ3OE was used as a probe to identify clones containing sequences coding
CC	for TCS in a T.kirilowii genomic library. Five clones were eventually
CC	isolated and sequenced, including clone 12. They were found to have
CC	homology to the alpha-TCS coding sequence. See also AAQ06343-Q06350
XX	
SQ	Sequence 1185 BP; 372 A; 249 C; 221 G; 343 T; 0 U; 0 Other;
 Alignment Scores:	
Pred. No.:	5.5e-132 Length: 1185
Score:	1326.00 Matches: 267
Percent Similarity:	95.16% Conservative: 8
Best Local Similarity:	92.39% Mismatches: 14
Query Match:	92.60% Indels: 0
DB:	Gaps: 0
 US-09-905-247-1 (1-289) x AAQ06351 (1-1185)	
QY	1 MetlleArgPheLeuValLeuSerLeuLeuLeuThrLeuPheLeuThrProAla 20
Db	269 ATGATCAGATTCTTAGTCTTCCTTTGGTAATTCACCCCTCTTCCTAACAGCTCGTGCT 328
QY	21 ValGlUGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrlGlyValPhe 40
Db	329 GTGAGGGCGATGTAGCTTCGGTTTATCAGGTGCACAAGCAGTTCCCTATGAGTTTTTC 388
QY	41 lleSerAsnLeuArcGlysAlaLeuProAsnGluAArgLysLeuTyrrAspIleProLeuLeu 60
Db	389 ATTCCAATATGAGAAAAGCTCTTCCATATGAAGAAGAGACTATACCATATCTCTGTGA 448
QY	61 ArgSerSerLeuProGlySerGlnArgTyrlAlaLeuLeuIleHisLeuThrAsnTyrlAlaAsp 80
Db	449 CGTTCACCTCTTCAGGTTCCTCAAGCTACGATGATCATCTCACAAATATACGCCGAT 508
QY	81 GluThrIleSerValAlalaileAspValThrAsnValTyrlMetGlyTyrrArgAlaGly 100
Db	509 GAACACCATTTGAGTGGCCATAGACGTAAACGACGCTATGTTATGGATATCGCGTGGT 568
QY	101 AspThrSerTyrlPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrrValPheLys 120
Db	569 GATACATCTCATTTTTTCAACGAGGCTTCGCACTGAAAGCTGC AAAATATGTATTCAA 628
QY	121 AspAlaMetArgLysValThrLeuProTyrrSerGlyAsnTyrlGluArgLeuGlnThrAla 140
Db	629 GACGCTCAGCGAAAAAGTTACGCTCCATATTCGCAATATCAAAAGGCTTCAAATTCGA 688
QY	141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
Db	689 GCAGGCAAAATAAGGAAAAATATTCGGCTTGACCTACCTGCTTGACAGTGCCATTACC 748
QY	161 ThrLeuPheTyrlTyrrAsnAlaAsnSerAlalaserAlaLeuMetValLeulleGlnSer 180
Db	749 AANTTGHTTTACTACAGCCCAATTCGCTCGCTGGCAGCTATATGGTACTATTTCAGTCG 808
QY	181 ThrSerGluAlaAlaArgTyrlLyapheilleGluGlnGlinilleGlyVsArgValAspLys 200
Db	809 ATGTCCTGAGGCTCGAGGTATAAATTTATTCAGCNACAAATTTGGAGCGGTGTGTACAAA 868
QY	201 ThrPheLeuProSerLeuAlaalleilerLeuGluAsnSerTyrrPseSerAlaLeuSerLys 220
Db	869 ACCTTCCTACCAAGTCTTGCAATATAAGTTTGGAAAATAGTTGCTGCTCTCTCAAG 928
QY	221 GlnIleGlnIleAlaSerThrAsnAsnClyGlnPheGluSerProValValLeuIleAsn 240

RESULT 15
AAQ06346
ID AAQ0
XX
AC AAQ0
XX
DT 06-F
XX

Search completed: September 14, 2004, 20:19:35
Job time : 393 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 14, 2004, 19:57:42 ; Search time 2476 Seconds
(without alignments)
3485.529 Million cell updates/sec

Title: US-09-905-247-1
Perfect score: 1432
Sequence: 1 MIFLVLSLILTLFTTPA.....AAMDDVPMTQSGCGSYAL 289

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09905247/runat_10092004_103934_4726/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HRAFSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09905247@cgn_1_1_5180 @runat_10092004_103934_4726 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:**
2: em_esthum:**
3: em_estin:**
4: em_estmu:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_hic:**
9: gb_est1:**
10: gb_est2:**
11: gb_hic:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: em_gss_hum:**
18: em_gss_inv:**
19: em_gss_pln:**
20: em_gss_vrt:**
21: em_gss_fun:**
22: em_gss_man:**
23: em_gss_mus:**
24: em_gss_pro:**
25: em_gss_rod:**
26: em_gss_phg:**
27: em_gss_vrl:**
28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	230.5	16.1	665	14	CF227084	CF227084 IHSLEP22 S
2	222	15.5	661	14	CF227046	CF227046 IH202 sub
3	220	15.4	837	9	AW053634	AW053634 L30-1401T
4	220	15.4	1021	10	BE036639	BE036639 MP03B03 M
5	214	14.9	662	14	CF227047	CF227047 IH215 sub
6	198.5	13.9	658	13	BQ583480	BQ583480 E011979-0
7	197.5	13.8	601	14	CA198032	CA198032 SCEZAD107
8	193	13.5	1033	10	BE036541	BE036541 MP01B07 M
9	181	12.6	420	14	CF227009	CF227009 IH149 sub
10	181	12.6	993	10	BE035039	BE035039 MM02A03 M
11	177.5	12.4	1038	10	BE033546	BE033546 MP03A09 M
12	171	11.9	698	14	CA838926	CA838926 MCT021H09
13	169.5	11.8	667	14	CA838446	CA838446 MCT016F04
14	169.5	11.8	678	14	CA840373	CA840373 MCT036G03
15	169.5	11.8	811	10	BE034055	BE034055 MG04C05 M
16	166.5	11.6	644	14	CA838401	CA838401 MCT016B04
17	163.5	11.4	611	13	BQ490182	BQ490182 34-E9435-
18	156.5	10.9	661	14	CA833333	CA833333 MCS038A11
19	156.5	10.9	689	14	CA833333	CA833333 MCT026C11
20	156.5	10.9	689	14	CA833383	CA833383 MCS021A09
21	152.5	10.6	649	14	CA838757	CA838757 MCT020B02
22	152.5	10.6	649	14	CA833330	CA833330 MCT026C11
23	152.5	10.6	671	14	CA839511	CA839511 MCT028D04
24	150.5	10.5	639	14	CA835855	CA835855 MCS004G12
25	143.5	10.0	405	14	T24255	T24255 CFS1345 lam
26	140.5	9.8	1094	11	AY105813	AY105813 Zea mays
27	139.5	9.7	816	10	BE035038	BE035038 MM02A01 M
28	132	9.2	639	10	BE037217	BE037217 MP18B02 M
29	126.5	8.8	677	14	CA174183	CA174183 SCJFST100
30	126	8.8	560	14	CF974192	CF974192 PSU meup2
31	126	8.8	607	14	CA222363	CA222363 SCSGFL403
32	125	8.7	680	14	CA198271	CA198271 SCCCF1300
33	124	8.7	419	14	CF226983	CF226983 IH114 sub
34	124	8.7	422	14	CF227044	CF227044 IH198 sub
35	123	8.6	620	13	CA080068	CA080068 SCSGAM109
36	121.5	8.5	672	9	AW053585	AW053585 L30-2143T
37	121.5	8.5	680	13	CA076205	CA076205 SMCAM110
38	121	8.4	765	14	CA195191	CA195191 SCEZSB109
39	120.5	8.4	730	13	CA088699	CA088699 SCRFAM212
40	118.5	8.3	498	13	BQ588134	BQ588134 E012337-0
41	118.5	8.3	603	13	BQ590856	BQ590856 E012599-0
42	118.5	8.3	660	14	CK113535	CK113535 UR108TF08
43	118.5	8.3	752	13	CA122188	CA122188 SCFUR107
44	118	8.2	713	12	BG367523	BG367523 HVSME1001
45	118	8.2	817	12	BG369108	BG369108 HVSME1002

ALIGNMENTS

RESULT 1
LOCUS CF227084
DEFINITION IHSLEP22 subtracted Iris tepal cDNA library Iris hollandica CDNA
ACCESSION CF227084
VERSION CF227084.1 GI:33438122
KEYWORDS EST.
SOURCE Iris hollandica
ORGANISM Iris hollandica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
Iris.

1 (bases 1 to 665)
van Doorn,W.G., Balk,P.A., van Houwelingen,A.M., Hoeberichts,F.A.,
Hall,R.D., Vorst,O., van der Schoot,C. and van Wordragen,M.F.
Gene expression during anthesis and senescence in Iris flowers
Unpublished (2003)
Contact: Wouter G. van Doorn
Institute for Agrotechnological Research (ATO)
Wageningen University and Research Centre
P.O. Box 17, 6700 AA, Wageningen, The Netherlands
Tel: 0031 317 475105
Fax: 0031 317 475347
Email: wouter.vandoorn@wur.nl.

FEATURES
source
1..665
/organism="Iris hollandica"
/mol_type="mRNA"
/cultivar="Blue Magic"
/db_xref="taxon:35876"
/clone="IHSUP22"
/tissue_type="sepals"
/clone_lib="subtracted Iris tepal cDNA library"
/notes="Vector: pGEMT-Easy; mRNA isolated from Iris sepals
was subtracted using pedicel material as the driver. Equal
amounts of mRNA from each developmental stage were pooled
for cDNA synthesis. The pGEM-T Easy Vector System
(Promega) was used to construct the cDNA library."

ORIGIN
Alignment Scores: 665
Pred. No.: 71
Score: 230.50
Percent Similarity: 51.98%
Best Local Similarity: 31.28%
Query Match: 16.10%
Indels: 24
Gaps: 7

US-09-905-247-1 (1-289) x CF227084 (1-665)

QY 36 SerTyrGlyValPheIleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyr 55
Db 6 ACCTATAGTGGTTCCTACAGACCTTAAAGCGGCTTAAGCAGTGAAGTAGCGTGCAT 65
QY 56 AsplleProLeuLeuArgSerSerLeuProGlySerGlnArgTyrAlaLeuHisLeu 75
Db 66 GATATTCTCTTGGTTCCTCGCAATCCGGTCTCAGCAGGATTTGCTCTTTGGCGATC 125
QY 76 ThrAsnTyrAlaAspGluThrIleSerValAlaIleAspValThrAsnValTyrIleMet 95
Db 126 TTTGACTGGGCCAACAGCCGCGATCAGCTAGTGTAAATCGCGTCAATGCTTATGGTG 185
QY 96 GlyTyrArgAlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAla 115
Db 186 GCTTATCAAGCTCAAAATGGTTTACTTGTCTTCCGCGC-----ACTCCGCCCAAT 236
QY 116 LysTyrValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGlu 135
Db 237 CCTCAGGTTTATGGAAACAAACCCACCGCCCTTACT-----TTACAGGTAGTATGGT 290
QY 136 ArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeu 155
Db 291 GCATTCAAAATGTCGGAACCAAGTAGAGAAATATTTGATGGGAATCAATCCACTT 350
QY 156 AspSerAlaIleThrThrLeuPheTyrTyrAsnAlaAsn-----Ser 169
Db 351 GCAACGGCAATCTCGACACTC-----CACAACTGCTCTCCGCCCTACGGTGGAAACATCC 404
QY 170 AlaAlaSerAlaLeuMetValIleGlnSerThrSerGluAlaAlaArgTyrLysPhe 189
Db 405 GTGGCTGTTCCCTTATCGTTTGTATCCAAATTTGGTCTCGAGACCGACGATTCAGACG 464
QY 190 IleGlnGlnIleGlyLysArg---ValAspLys-----Thr 201
Db 465 ATGAGCAGACGATTCAACAAACATCATAGACAGGTCACCCCAATCCGCTACGACAA 524

202 PheLeuProSerLeuAlaIleSerLeuGluAsnSerTrpSerAlaLeuSerLysGln 221
Db 525 TTTCCGCCCTAGTGTGGCAATATAGACCTCAGCACTACTG-CAAAACCCCTCCACTGAA 583
QY 222 IleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuAlaAsnAla 241
Db 584 GTTCAGCGAGCC-----GAAGGAGTAGGTTTTCGAGCCAGNTAAATTGCAGGNTTCC 637
QY 242 GlnAsnGlnArgValThrIle 248
Db 638 GTCCAGCAACACAGTCGGTATC 658

RESULT 2
CF227046 661 bp mRNA linear EST 04-AUG-2003
LOCUS
DEFINITION
IH202 subtracted Iris tepal cDNA library Iris hollandica cDNA clone
IH202, mRNA sequence.
ACCESSION
VERSION
CF227046.1 GI:33438082
KEYWORDS
EST.
SOURCE
Iris hollandica
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
Iris.
REFERENCE
1 (bases 1 to 661)
van Doorn,W.G., Balk,P.A., van Houwelingen,A.M., Hoeberichts,F.A.,
Hall,R.D., Vorst,O., van der Schoot,C. and van Wordragen,M.F.
Gene expression during anthesis and senescence in Iris flowers
Unpublished (2003)
Contact: Wouter G. van Doorn
Institute for Agrotechnological Research (ATO)
Wageningen University and Research Centre
P.O. Box 17, 6700 AA, Wageningen, The Netherlands
Tel: 0031 317 475105
Fax: 0031 317 475347
Email: wouter.vandoorn@wur.nl.

FEATURES
source
1..661
/organism="Iris hollandica"
/mol_type="mRNA"
/cultivar="Blue Magic"
/db_xref="taxon:35876"
/clone="IH202"
/tissue_type="sepals"
/clone_lib="subtracted Iris tepal cDNA library"
/note="Vector: pGEMT-Easy; mRNA isolated from Iris sepals
was subtracted using pedicel material as the driver. Equal
amounts of mRNA from each developmental stage were pooled
for cDNA synthesis. The pGEM-T Easy Vector System
(Promega) was used to construct the cDNA library."

ORIGIN
Alignment Scores: 661
Pred. No.: 71
Score: 222.00
Percent Similarity: 50.67%
Best Local Similarity: 31.56%
Query Match: 15.50%
Indels: 24
Gaps: 7

US-09-905-247-1 (1-289) x CF227046 (1-661)

QY 1 MetIleArgPheLeuValLeuSer-----LeuLeuIleLeuThrLeuPheLeuThr 17
Db 7 ATGAAGACGTGTTGATCTTGGCAGTGCAGTGCATTTTATGGCTAGCTTGGACCT 66
QY 18 ThrProAlaValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerTyr 37
Db 67 GTAGCTTCCATTGAA---ACAGTCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
QY 38 GlyValPheIleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIle 57

```

Db      124 AGTGGTTTCCTACAGACCCCTAGAACGCTTAAAGCAGTGGAGTAGCGTGCATGATATT 183
QY      58 ProLeuLeuArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuLeuLeuLeuLeu 77
Db      184 CTTTGTGCTGGCGCAATCCGGTCCCGAGAGGATTGCTCTTGTGCAACTCTTTGAT 243
QY      78 TyrAlaAspGluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyr 97
Db      244 TGGGCAACACCGCCGATCAGCTGGTCTTAATCGGTCAATGCTTATCTGTAGCTTAT 303
QY      98 ArgAlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyr 117
Db      304 CAAGCTAAATTCGTATTACTGTGTT-----TCCGACACTCCCGCCAATCTCAG 354
QY      118 ValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeu 137
Db      355 CTTTATGGAAGCAACCCGACCGC-----CTTAGTTTACAGGAGCTACATGCTCTT 408
QY      138 GlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSer 157
Db      409 CAAATGTGTGGCAAAACAAGTAGAAAAATATTGATCTGGGAATCAATCCACTTCGAACG 468
QY      158 AlaIleThrThrLeuPheTyr-----TyrAsnAlaAsnSerAlaAlaSerAla 173
Db      469 GCATCAGCACACTCCA-CAACTGGGCTCCGCTACGCTGAGGAGACATCCGTGGCTCGTCC 527
QY      174 LeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGln 193
Db      528 CTTATCGTCTTGTCACTCGTCTCCGAGCAGCAGCATTCAGAGCGATCGAGCAACGA 587
QY      194 IleGlyLys-----ArgValAspLysThrPheLeuPro 204
Db      588 GTTACGAGCAGCATATAGACGCGCTACCCCTATCCGTTACGAC---AGCTTTCGTCCT 644
QY      205 SerLeuAlaIleIle 209
Db      645 GGTGTGGCATAATA 659

RESULT 3
AW053634
LOCUS
DEFINITION
L30-140173 Ice plant Lambda Uni-Zap XR expression library, 30 hours
NaCl treatment Mesembryanthemum crystallinum cDNA clone L30-1401
5', mRNA sequence.
AW053634
AW053634.1 GI:5916827
EST.
Mesembryanthemum crystallinum (common iceplant)
Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Alceaceae; Mesembryanthemum.
1 (bases 1 to 837)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
Poly(A) tail, 10 nt: 838. .847
PCR Primers
FORWARD: T7
BACKWARD: T3
Insert Length: 847 Std Error: 5.00
Plate: L30-15 row: A column: 5
Seq primer: T3
High quality sequence stop: 450.

FEATURES
Source
Location/Qualifiers
1..837
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
/clone="L30-1401"
/tissue_type="Leaf, 30 h 0.4M NaCl"
/dev_stage="Six week old"
/clone_lib="Ice plant Lambda Uni-Zap XR expression
library, 30 hours NaCl treatment"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 1,3e-16 Length: 837
Score: 220.00 Matches: 69
Percent Similarity: 51.20% Conservative: 38
Best Local Similarity: 33.01% Mismatches: 73
Query Match: 15.36% Indels: 30
DB: 9 Gaps: 8

US-09-905-247-1 (1-289) x AW053634 (1-837)
QY 35 SerSerTyrGlyValPheIleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeu 54
Db 5 AACACATATTCACATTCATGACATCTCTCGAGCTCAACTCTCTGGCACCAACCGCATGT 64
QY 55 TyrAspIleProLeuLeuArgSerSerLeuProGlySerGlnArgTyrAlaLeuLeuLeu 74
Db 65 ---CAATCCCGCGACACGGTCAACGCCCGGATCTCCAAACGATTCGTCTCGTCGAC 121
QY 75 LeuThrAsnTyrAlaAspGluThrIleSerValAlaIleAspValThrAsnValTyrIle 94
Db 122 CTCAAAAACAACCTCGCAAAAGACCATCACATCTCGCAATCGACGTGACAGCGTCTATGTC 181
QY 95 MetGlyTyrArg-----AlaGlyAspThrSerTyrPhePheAsnGlyAlaSer 110
Db 182 GTGGCTTACCGCGACAAGCTTGGCGGAAAGACGCTGCCCAACTCTTAGCGATGCT--- 238
QY 111 AlaThrGluAlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyr 130
Db 239 CCCACGGTCTAGGAAACACCTCTTCAAGGGCGGAGGTTCCG---AACATTGCGTTC 295
QY 131 SerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeu 150
Db 296 GGAGGGAGTTACATAGCCTTAGAGAGGCTGCAAGCAAGCGCGAATGCAATAGACTTA 355
QY 151 GlyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe-----TyrTyrAsnAla 167
Db 356 GGGGTGAATAACTAGAGTTTGGATCGAGTCGGTCTTTTGTGAAGAACCCGATTATATGGG 415
QY 168 AsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyr 187
Db 416 CAGGTTGAGGCCAAATCTTCTGTGATGCTATACAGATGGTTCTTGAAGCAGCAGCGTTT 475
QY 188 LysPheIleGluGln-----GlnIleGlyLysArgValAspLysThrPheLeuPro 204
Db 476 AAGTATATTGAAGTAAGGGGACCCCAAGTGGG-----TTACATGGTGGTGGTCAACACC 529
QY 205 SerLeuAlaIleIleSerLeuGluAsnSerTrpSerAlaLeuSerLysGlnIleGlnIle 224
Db 530 GACCC- GAAGTCTGAGTTTGGAGAAC-----
QY 225 AlaSerThrAsnAsnGlnPheGlu 233
Db 556 -----AATGGGGAAGATTCAG 573

RESULT 4
BE036639
LOCUS
DEFINITION
MP03B03 MP Mesembryanthemum crystallinum cDNA 5' similar to
ribosome-inactivating protein gelonin, mRNA sequence.

```

1021 bp mRNA linear EST 07-JUN-2000
 BE036639
 LOCUS
 DEFINITION
 High quality sequence stop: 450.

```

ACCESSION BE036639
VERSION BE036639.1 GI:8331648
KEYWORDS EST.
SOURCE Mesembryanthemum crystallinum (common iceplant)
ORGANISM Mesembryanthemum crystallinum
REFERENCE Bohnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M.,
AUTHORS Ferreira,H.J., Kawasaki,S., McColough,A., Michalowski,C.B.,
Palacio,C., Scara,G., Wheeler,M. and Zepeda,G.R.
TITLE Functional Genomics of Plant Stress Tolerance
JOURNAL Unpublished (2000)
COMMENT Contact: Michalowski,C.B.
University of Arizona
Bio Sciences West room 513, Tucson, AZ 85721, USA
Tel: 520-621-7982
Fax: 520-621-1697
Email: cbm@u.arizona.edu
An open reading frame exists.
Insert Length: 1 Std Error: 0.00.
FEATURES
Location/Qualifiers
1..1021
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
/tissue_type="apical meristem and leaf primordia"
/dev stage="6 weeks"
/clone lib="MP"
/note="3 d 500mM NaCl"
source
1..1021
Alignment Scores:
Pred. No.: 1.79e-16 Length: 1021
Score: 220.00 Matches: 72
Percent Similarity: 54.03% Conservative: 42
Best Local Similarity: 34.12% Mismatches: 75
Query Match: 15.36% Indels: 23
DB: 10 Gaps: 8
US-09-905-247-1 (1-289) x BE036639 (1-1021)
QY 25 ValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheIleSerAsnLeu 44
DB 225 GTGACATTCACCTC-----GCCAACCCACACATATTCACATTCATGACATCTCTC 278
QY 45 ArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSerLeu 64
DB 279 CGAGTCACCTCTCTGGCACACCGCATGT---CAAAATCCCGTCGACACGGTCAACCGCA 335
QY 65 ProGlySerGlnArgTyrAlaLeuLeuIleHisLeuThrAsnTyrAlaAspGluThrIleSer 84
DB 336 CGGGATCTCCACAGTATCTCTCTGCTGCTGACCTCAAAACACCTGCAAAAGACCATCACA 395
QY 85 ValAlaIleAspValThrAsnValTyrIleMetGlyTyrArg-----AlaGly 100
DB 396 CTCGCAATCGAGTCGACACAGCTGATGTCGTGGCTACCGGCACAGCTTGGCGGAAA 455
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
DB 456 GACCGTGCCAACTCTCTTAGCATGTCT---CCACGGTCTGCTAGGAACAACTCTTCAAG 512
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
DB 513 GCGCGCACCGTTCTGG---AACATTCGCTTCGGAGGAGTATACATCCCTTAGAGAGGCT 569
QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
DB 570 GAAACCAAGCGGCAATGCAATAGATAGAGTGGGTGATAACTAGAGTTTGCATCGAG 629
QY 161 ThrLeuPheTyr-----TyrAsnAlaSerAlaLeuMetValLeu 177
DB 177

```

```

Db 630 TCGGT-TTTTGTGAGAACCCGATTATATGGCATGTTGAGCCAAATCTTGCTGATTGCT 688
QY 178 IleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnIleGlyLys--- 196
Db 689 ATACAGATGGTTTCTGACGACGACCGTGTAAAGTATATTGAAAGTAGGTGACCCAAAGT 748
QY 197 -----ArgValAspLysThrPheLeuProSerLeuAlaIleIleSerLeu-GI 212
Db 749 GGTATACATGGCTCGTTCAAC-----CCGACCCGAAAGTGTGAGTTTGGGA 796
QY 212 uAsnSerTyrSerAlaLeuSerLysGlnIle 222
Db 797 GAACAATTTGGGAGAGATTTTCGCAAGAAATTA 827
RESULT 5
CF227047 662 bp mRNA linear EST 04-AUG-2003
LOCUS IH215 subtracted Iris tepal cDNA library Iris hollandica cDNA clone
DEFINITION IH215, mRNA sequence.
ACCESSION CF227047.1 GI:33438083
VERSION CF227047
KEYWORDS EST.
SOURCE Iris hollandica
ORGANISM Iris hollandica
REFERENCE 1 (bases 1 to 662)
AUTHORS van Doorn,W.G., Balk,P.A., van Houwelingen,A.M., Hoerberichts,F.A.,
Hall,R.D., Vorst,O., van der Schoot,C. and van Woudagen,M.F.
TITLE Gene expression during anthesis and senescence in Iris flowers
JOURNAL Unpublished (2003)
COMMENT Contact: Wouter G. van Doorn
Institute for Agrotechnological Research (ATO)
Wageningen University and Research Centre
P.O. Box 17, 6700 AA, Wageningen, The Netherlands
Tel: 0031 317 475105
Fax: 0031 317 475347
Email: wouter.vandoorn@wur.nl.
FEATURES
Location/Qualifiers
1..662
/organism="Iris hollandica"
/mol_type="mRNA"
/cultivar="Blue Magic"
/db_xref="taxon:35876"
/clone="IH215"
/tissue_type="sepals"
/clone_lib="subtracted Iris tepal cDNA library"
/note="vector: pGEMT-Easy; mRNA isolated from Iris sepals
was subtracted using pedicel material as the driver. Equal
amounts of mRNA from each developmental stage were pooled
for cDNA synthesis. The pGEM-T Easy Vector System
(Promega) was used to construct the cDNA library."
ORIGIN
Alignment Scores:
Pred. No.: 5.07e-16 Length: 662
Score: 214.00 Matches: 63
Percent Similarity: 52.74% Conservative: 43
Best Local Similarity: 31.34% Mismatches: 82
Query Match: 14.94% Indels: 14
DB: 14 Gaps: 5
US-09-905-247-1 (1-289) x CF227047 (1-662)
QY 1 MetIleArgPheLeuValLeuSer-----LeuLeuIleLeuThrPheLeuThr 17
Db 40 ATGAGACGCTGGTTGATCTTGGCAGTGAGCTGATTTATGGCTACGATTCTGCACCT 99
QY 18 ThrProAlaValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerTyr 37
Db 100 GTAGCTTCCATTGAA---ACAGTCGAGTTCCGTGTCCTGCTGGGACCAACGACCTAT 156

```

```

QY 38 GlyValPheIleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIle 57
Db 157 AGTGCTTTCTCAGACCGCTGAGCAAGCTGAAGTAGTGAAGTGCATGATATT 216
QY 58 ProLeuLeuArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsn 77
Db 217 CTTTGTCTGCTGGGCAATCCGGTCCAGCAGGATTGCTCTCTGTGCAACTCTTTGAT 276
QY 78 TyrAlaAspGluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyr 97
Db 277 TGGGACAACACCGCATCAGCTGGTGTCTTAATCGCGTCAATGCTTATCTGTAGCTTAT 336
QY 98 ArgAlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaLysTyr 117
Db 337 CAAGCTAAAAAATCGTTATTACTTGTCTT-----TCGACACTCCCGCCAATCCTCAG 387
QY 118 ValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeu 137
Db 388 CTTTATGGAGCAACCCGACCGC-----CTTAGTTTACAGGCGAGCTPACATTCACCTT 441
QY 138 GlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSer 157
Db 442 CAAAATGTTGCAAAACAAAGTAGAGAAATATTGATCTGGGAATCAATCCACTTGCACG 501
QY 158 AlaIleThrLeuPheTyrTyrAsnAla-----AsnSerAlaAlaSerAla 173
Db 502 GCAATCAGACACTCCAACTCCAGTGGC-TCCGCTACGGTGGAGACATCCGTGGCTCGTTCC 560
QY 174 LeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnGln 193
Db 561 CTTATCGTCTTGAATCAACTGCTCCGAGACCGCAGCATTCAGAGCGATCGACACACGAG 620
QY 194 Ile 194
Db 621 TTA 623

RESULT 6
BQ583480
LOCUS
DEFINITION
E011979-024-005-N05-SP6 MP1Z-ADIS-024-inflorescence Beta vulgaris
CDNA clone 024-005-N05 5-PRIME, mRNA sequence.
ACCESSION
BQ583480
VERSION
BQ583480.1 GI:26113057
KEYWORDS
EST.
SOURCE
Beta vulgaris
ORGANISM
Beta vulgaris
REFERENCE
1 (bases 1 to 658)
Herwig,R., Schulz,B., Weisshaar,B., Hennig,S., Steinfath,M.,
Drungowski,M., Stahl,D., Wruck,M., Menze,A., O'Brien,J., Lehrach,H.
and Radelof,U.
Caryophyllales; Amaranthaceae; Beta.
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Construction of a 'unigene' cDNA clone set by oligonucleotide
fingerprinting allows access to 25 000 potential sugar beet genes
Plant J. 32 (5), 845-857 (2002)
22362189
12472698
COMMENT
Contact: Weisshaar B
ADIS DNA core facility at MP1Z
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@mpiz-koeln.mpg.de
Insert Length: 658 Std Error: 0.00
Plate: 5 row: N column: 05
Seq primer: SP6; CATACATTAGGTGACACTATAG.
Location/Qualifiers
1. .658
/organism="Beta vulgaris"
/mol type="mRNA"
/cultivar="KWS2320 (double haploid, monogerm breeding
line)"

```

```

/db_xref="GABI:183219"
/db_xref="taxon:161934"
/clone="024-005-N05"
/tissue_type="inflorescence"
/lab_host="EMDH10B"
/clone_lib="MP1Z-ADIS-024-inflorescence"
/note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;
cDNA library from sugar beet, library provided by KWS
Kleinwanzlebener Saatzaucht AG Einbeck, Germany, contact:
b.schulz@kws.de; cloning sites SalI-NotI, primer sites and
orientation:
SP6-Sali-CCACGGCTCCG-5prime-cDNA-polyA-CC-NotI-T7; Note:
Sequencing granted in the context of the GABI-Beet
project, local PI: Dr. Katharina Schneider, coordinator:
Prof. Christian Jung; Sequence submission managed by
RZPD/GABI-Primary database: http://gabi.rzpd.de"

ORIGIN

Alignment Scores:
Pred. No.: 4,56e-14 Length: 658
Score: 198.50 Matches: 57
Percent Similarity: 49.70% Conservative: 26
Best Local Similarity: 34.13% Mismatches: 75
Query Match: 13.86% Indels: 9
DB: 13 Gaps: 5

US-09-905-247-1 (1-289) x BQ583480 (1-658)

QY 26 SerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeuArg 45
Db 166 ACATACGACCTGAGCGGGGTACA--AAAGCTACTCTCAATCTTGACAAATTATGCG 222
QY 46 LysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSerLeuPro 65
Db 223 AAAAAATTG---GAAGCAACAAGGCTTGCACATGCCAGTCGACGAGCAAACTTACCA 279
QY 66 GlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThrIleSerVal 85
Db 280 GCAGCGCGGAATATGTTTGTAGTGATCTCAAGTTACCCAGTACAAATATGTCACCTTA 339
QY 86 AlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAspThrSerTyr--- 104
Db 340 GCTATAGAGTTAATAATGTTAATCTGTGGTCTTATCGTGTGGCGGCAAACTCAAGGC 399
QY 105 ---PhePheAsnGlyAlaSerAlaThrGluAlaLysTyrValPheLysAspAlaMet 123
Db 400 ACATTCTCTTAATGAGCCAGATGTTACATTTGCCAGGTCTAAGCTTTTGTATAAGATGAT 459
QY 124 ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143
Db 460 AATCAACCAACACTTCGATATGACAGCAACTACGCTAATATAGAG--GCGGTGGTGGC 516
QY 144 IleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe 163
Db 517 TCCAGGAAGGACCTTCTCTGGGCTTCAGGAATTACGGAAGAGAAATTGAGCTCTTT 576
QY 164 TyrTyrAsn-----AlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGln 179
Db 577 GACCAAAATAAACCAAGTGTGAGCGGAGGAGCTCAATTCTTATCAAGCAATTCAA 636
QY 180 SerThrSerGluAlaAlaArg 186
Db 637 ATGTTATCTGAGGCTGCGCA 657

RESULT 7
CA198032
LOCUS
DEFINITION
SCEAD1078B05.g AD1 Saccharum officinarum cDNA clone SCEAD1078B05
S, mRNA sequence.
ACCESSION
CA198032
VERSION
CA198032.1 GI:35228882
KEYWORDS
EST.
SOURCE
Saccharum officinarum

```

Db 297 GAGTCGTACACCTCCACAG-----TCTTCATTTCGACCTAGGCAACGATGATTGGT 350

QY 211 LeuGluAsnSerTyrSerAlaLeuSerLysGlnLeuAlaSerThrAsnAsnGly 230

Db 351 TTTGAGAACAACTGGGAGACCTCTCTGACAGATCAAACTAGCGAGGTA-----TTC 404

QY 231 GlnPheGluSerProValValLeuLeuAlaGlnAsnGlnArgValThrLeuThrAsn 250

Db 405 GCCTTCCAACTGCT-----ATTAACTTGGCGGAA-----GTTGTGACAAT 449

QY 251 ValAspAlaGlyValValThrSerAsnLeuLeuLeu 263

Db 450 GTTCGATCAGCAGTTATTAAGGCTTTGTTCTTAATGCTG 488

RESULT 8

BE036541 1033 bp mRNA linear EST 07-JUN-2000

LOCUS MP01807 MP Mesembryanthemum crystallinum cDNA 5' similar to

DEFINITION antiviral protein, mRNA sequence.

ACCESSION BE036541

VERSION BE036541.1 GI:8331550

KEYWORDS EST.

SOURCE Mesembryanthemum crystallinum (common iceplant)

ORGANISM Mesembryanthemum crystallinum

REFERENCE 1 (bases 1 to 1033)

AUTHORS Bohnert, H. J., Borchert, C., Brazille, S., Brooks, J., Eaton, M., Ferreira, H., Kawasaki, S., McColough, A., Michalowski, C. B., Palacio, C., Scara, G., Wheeler, M. and Zepeda, G. R.

TITLE Functional Genomics of Plant Stress Tolerance

JOURNAL Unpublished (2000)

COMMENT Contact: Michalowski, C. B. University of Arizona Bio Sciences West room 513, Tucson, AZ 85721, USA Tel: 520-621-7982 Fax: 520-621-1697 Email: cbm@u.arizona.edu An open reading frame exists.

FEATURES

source

1..1033

/organism="Mesembryanthemum crystallinum"

/mol_type="mRNA"

/db_xref="taxon:3544"

/tissue_type="apical meristem and leaf primordia"

/dev_stage="6 weeks"

/clone_lib="MP"

/note="3 d 500mM NaCl"

ORIGIN

Alignment Scores:

Pred. No.: 4,72e-13 Length: 1033

Score: 193.00 Matches: 62

Percent Similarity: 47.11% Conservative: 44

Best Local Similarity: 27.56% Mismatches: 73

Query Match: 13.48% Indels: 46

DB: 10 Gaps: 10

US-09-905-247-1 (1-289) x BE036541 (1-1033)

QY 88 AspValThrAsnValTyrIleMetGlyTyrArg-----AlaGlyAspThrSer 103

Db 38 GAGGCGCAAGAGGCTTATGTCGTGCTACCGCAAGCTTGGCGGAAAGACCGTGCC 97

QY 104 TyrPhePheAsnGlyValSerAlaThrGluAlaAlaLysTyrValPheLysAspAlaMet 123

Db 98 AACTTCCTTAGGATGCT---CCCAGCTGCTAGAACAACTCTTCAAGGCGGAGC 154

QY 124 ArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLys 143

Db 155 GTTCGG---AACATTGCGTTCGAGGAGTTACATAGCTCTAGAGAAGGCTGCAAGCAA 211

ORGANISM Saccharum officinarum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.

AUTHORS 1 (bases 1 to 601)

TITLE Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.

JOURNAL The libraries that made SUCEST

COMMENT Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: Clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bccccenter.fcav.unesp.br>

Plate: 078 row: B column: 05

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1..601

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCEZAD1078B05"

/lab_host="DH10B"

/clone_lib="AD1"

/note="Organ: seedlings inoculated with Gluconacetobacter diazotrophicans; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [seedlings inoculated with Gluconacetobacter diazotrophicans]. cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a Sepharose CL-2B 40cm-column and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:

Pred. No.: 5,27e-14 Length: 601

Score: 197.50 Matches: 58

Percent Similarity: 51.45% Conservative: 31

Best Local Similarity: 33.53% Mismatches: 73

Query Match: 13.79% Indels: 11

DB: 14 Gaps: 6

US-09-905-247-1 (1-289) x CAL98032 (1-601)

QY 91 AsnValTyrIleMetGlyTyrArgAlaGlyAspThrSerTyrPheAsnGlyAlaSer 110

Db 3 AATGTTTACATCATGTTGTTATCTTATGTTACACAGATATCTTCCACGGAGCTGAA 62

QY 111 AlaThrGluAlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyr 130

Db 63 ---GATGGGAGATCGGCTTCTTCACAGATGACGCCACCTTCAGCAACAGGAGA 119

QY 131 SerGlyAsnTyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeu 150

Db 120 TCAGCCAGCTATCGCAGATGAGAGCGGAGGACACATAAATAGGCTAACATCGACCTA 179

QY 151 GlyLeuProAlaLeuAspSerAlaIleThrThrLeuPheTyrTyrAsnAlaAsnSerAla 170

Db 180 GGTCTGCTAGTATGAGCGGCCCATAGAACCTTGTAGCTGAGGAGACAGAGATGTA 239

QY 171 AlaSerAlaLeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIle 190

Db 240 AGGTCA---CTTCTAGTTGTTATCCAAATGGTTGCAGAGGGTGTGGGTTTGGGAGATA 296

QY 191 GluGlnGlnIleGlyLysArgValAspLysThrPheLeuProSerLeuAlaIleSer 210

/cell_type="epidermal bladder cells"
 /dev_stage="12 weeks old"
 /clone_lib="WM"
 /notes="Vector: Bluescript SK+; Site_1: EcoRI; Site_2:
 XhoI; Plants stressed 6 weeks in 500mM NaCl"

ORIGIN

Alignment Scores: 1.45e-11 Length: 993
 Pred. No.: 181.00 Matches: 55
 Score: 51.76% Conservative: 33
 Percent Similarity: 32.35% Mismatches: 66
 Best Local Similarity: 12.64% Indels: 16
 Query Match: 10 Gaps: 7
 DB:

US-09-905-247-1 (1-289) x BE035039 (1-993)

QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaLysTyrValPheLys 120
 Db 787 GATCGAGCCAACTTCCTTAGGATGCT--CCTACAGTCGCTAAGAAACCTTCTTTAAG 731
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
 Db 730 GGAGCGACCGTTCGA--AACATTCATTTGGTGGTAACATACATAGCTCTCGGAAAGCT 674
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyValProAlaLeuAspSerAlaIleThr 160
 Db 673 GCCGGTCAAGACGAAACAGCATATTAGTTGGGGCTTGTAACTAGATTGCGATCGAG 614
 QY 161 ThrLeuPhe-----TyrTyrAsnAlaAsnSerAlaLysSerAlaLeuMetValLeu 177
 Db 613 TCGATTATGTTAATAAAACGATCGTGGAAACCTCGAGGCGCAATTCCTTACTGATTGCC 554
 QY 178 IleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArg 197
 Db 553 ATCCAGATGGTTTCAGAGACGACGGTTCATATATATCGAGCTTAAGTGAACGAAAGT 494
 QY 198 ValAspLys---ThrPheLeuProSerLeuAlaIleLeuSerLeuGluAsnSerTipSer 216
 Db 493 GGGTCACATGGTTCGTTCAAAACCGATCGGAAAGTATTAATCTGGAGAACAACTGGGGA 434
 QY 217 AlaLeuSerLysGlnIleGlnIleAlaSerThrAsn-----AsnGlyGlnPhe 232
 Db 433 AAGATTTCGGATGAGATTTCATAGTCAGTTCAGTTCGAAACACCATCTGTAATGTACGAAC 374
 QY 233 GluSerProValVal---LeuIleAsnAlaGlnAsn-----GlnArgVal 246
 Db 373 ATTTCCCTCGGATTAATCAATCTGATGCTGATGCTACACATGGAAGTGGATAAGATT 314
 QY 247 ThrIleThrAsnValAspAlaGlyValVal 256
 Db 313 GCCACTATAAACCTGACTTGGGGATCTC 284

BE033546 1038 bp mRNA linear EST 07-JUN-2000
 MF03A09 MF Mesembryanthemum crystallinum cDNA 5' similar to
 ribosome-inactivating protein, mRNA sequence.

ACCESSION BE033546
 VERSION BE033546
 KEYWORDS EST

SOURCE BE033546.1 GI:8328555

ORGANISM Mesembryanthemum crystallinum (common iceplant)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllales; Alzooaceae; Mesembryanthemum.

REFERENCE 1 (bases 1 to 1038)

AUTHORS Bohnert, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M.,
 Ferreira, H., Kawasaki, S., McColough, A., Michalowski, C.B.,
 Palacios, C., Scarpa, G., Wheeler, M. and Zepeda, G.R.

TITLE Functional Genomics of Plant Stress Tolerance

JOURNAL Unpublished (2000)

COMMENT Contact: Michalowski, C.B.

University of Arizona
 Bio Sciences West room 513, Tucson, AZ 85721, USA
 Tel: 520-621-7982
 Fax: 520-621-1697
 Email: cbm@u.arizona.edu
 An open reading frame exists.

FEATURES

Location/Qualifiers
 1..1038
 /organism="Mesembryanthemum crystallinum"
 /mol_type="mRNA"
 /db_xref="taxon:3544"
 /tissue_type="Root"
 /dev_stage="5-6 weeks old"
 /clone_lib="MF"
 /notes="Vector: Bluescript SK+; Site_1: EcoRI; Site_2:
 XhoI"

ORIGIN

Alignment Scores: 4.33e-11 Length: 1038
 Pred. No.: 177.50 Matches: 65
 Score: 49.50% Conservative: 34
 Percent Similarity: 32.50% Mismatches: 74
 Best Local Similarity: 12.40% Indels: 27
 Query Match: 10 Gaps: 8
 DB:

US-09-905-247-1 (1-289) x BE033546 (1-1038)

QY 56 AspIleProLeuLeuArgSerSerLeuProGly-SerGlnArgTyrAlaLeuIleHisLe 75
 Db 33 GATCNCGGCTCGCAGGAATTCGGTCGGTCGCGACCAACGATTCCTCTCGTGCACCT 92
 QY 75 uThrAsnTyrAlaAspGluThrIleSerValAlaIleAspValThr-AsnValTyrIleM 95
 Db 93 CAABAACAACCTCGCAAGACCATCACACTCGCAATCGCGTCGACAGAGCTCTATGTCG 152
 QY 95 etGlyTyrArg-----AlaGlyAspThrSerTyrPhePheAsnGlyAlaSerA 111
 Db 153 TGGCTACTCCGCGCAAGCTTGGCGGAAAGACCTGCGCACTTCTTACGATGCT---C 209
 QY 111 laThrGluAlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyrS 131
 Db 210 CCACGTCGCTAGGAACACCTCTTCAAGGCGCGACGGTTCGG---AACATGGTTCG 266
 QY 131 erGlyAsnTyrGluArgLeuGlnThrAlaLaglyLysIleArgGluAsnIleProLeuG 151
 Db 267 GAGGAGTTACATAGCTTTAGAGAAGGCTGCAAGAGCGGCAAGTCAATAGAGTTAG 326
 QY 151 lyLeuProAlaLeuAspSerAlaIleThrThrLeuPhe-----TyrTyrAsnAlaA 168
 Db 327 GGGTGAATAAACTAGAGTTCGCGATCGAGTCGGTTTTTGGTAAGAACCCGATTAATGGC 386
 QY 168 snSerAlaAlaSerAlaLeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrL 188
 Db 387 AGGTGAGCGCAAAATTCCTGCTGATGTATACAGATGTTCTTCAAGCAGCAGCGTTTA 446
 QY 188 yspHeileGluGlnIleGlyLysArg---ValAspLysThrPheLeuProSerLeuA 207
 Db 447 AGTTTATTGAAAGTAAAGTACCACCAAGTGGTGATGCTGTTGTTCAACCCGACCGCA 506
 QY 207 laIleIleSerLeuGluAsnSerTrp-----SerAlaL 218
 Db 507 AAGTCTGAGTTCGGTGATCAATTCGGTGAGCAATTTCCATGAGACTCATAGACAGCGC 566
 QY 218 euSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProVal 236
 Db 567 CGTGGAAACCT-----GCCTGACTAAT-----TTCCACGCGCTATT 604

RESULT 12

CA838926

LOCUS

DEFINITION

CA838926 698 bp mRNA linear EST 12-DEC-2002
 MCT021H09 170947 Ice plant Lambda Uni-Zap XR expression library, 5
 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase IV

Db	440	GAAGAGCGGACAAATGAAGACCGTGCACACTTCCCTTCAAGATGCATCT	---	ACGGTT	490
QY	114	AlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsn	133		
Db	497	GGCAGGAGGAGTCTCTTTCATCGGACACAGTTCGG--AACCTTGCGTTCAAGGGGACT	553		
QY	134	TyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuPro	153		
Db	554	TACACATCCTTAGAGATGCTGCAACACCAACGACGAGAACCATAGATAGTGGGTGGAT	613		
QY	154	AlaLeuAspSerAlaIleThrThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAla	173		
Db	614	AAACTAGAGTTGCGATCGATCGGTTTATGTAAGACA-----	652		
QY	174	LeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIle	190		
Db	653	-----TCGACAGTCAGAGAAATGAGGCCAAATCTTG	685		
RESULT 13					
LOCUS	CA838446				
DEFINITION	MCT016F04.166823 Ice plant Lambda Uni-Zap XR expression library, 5 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase IV (5:30 PM). Mesembryanthemum crystallinum cDNA clone MCT016F04 5, mRNA sequence.	667 bp	mRNA	linear	EST 12-DEC-2002
ACCESSION	CA838446				
VERSION	CA838446.1	GI:26566211			
KEYWORDS	EST.				
SOURCE	Mesembryanthemum crystallinum (common iceplant)				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Aizoaceae; Mesembryanthemum.				
REFERENCE	1 (bases 1 to 667)				
AUTHORS	Cushman, J.C.				
TITLE	An expressed sequence tag database for the common ice plant.				
JOURNAL	Mesembryanthemum crystallinum				
COMMENT	Unpublished (1997) Contact: Cushman JC Department of Biochemistry University of Nevada MS2000, Reno, NV 89557-0014, USA Tel: 775-784-1918 Fax: 775-784-1650 Email: jcushman@unr.edu PCR Primers FORWARD: T3 20mer BACKWARD: T7 21mer Plate: 016 row: F column: 04 Seq primer: T3 20mer High quality sequence stop: 667.				
FEATURES	Location/Qualifiers				
source	1..667				
	/organism="Mesembryanthemum crystallinum"				
	/mol_type="mRNA"				
	/db_xref="taxon:3544"				
	/clone="MCT016F04"				
	/tissue_type="leaf"				
	/dev_stage="five-week-old"				
	/clone_lib="Ice plant Lambda Uni-Zap XR expression library, 5 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase IV (5:30 PM)."				
	/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Library construction was performed according to Stratagene's recommended protocol for the Lambda UniZapXR vector and cDNA synthesis kit."				
ORIGIN					
Alignment Scores:					
Pred. No.:	2.16e-10	Length:	667		
Score:	169.50	Matches:	51		
Percent Similarity:	54.00%	Conservative:	30		
Best Local Similarity:	34.00%	Mismatches:	54		

```

/mol_type="mRNA"
/db_xref="taxon:3544"
/clone="MCT036G03"
/tissue_type="leaf"
/dev_stage="five-week-old"
/clone_lib="Ice plant Lambda Uni-Zap XR expression
library, 5 days 0.5 M NaCl treatment, Crassacean acid
metabolism, phase IV (5:30 PM)."
/notes="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; library construction was performed
according to Stratagene's recommended protocol for the
Lambda UniZapXR vector and cDNA synthesis kit."

```

ORIGIN

Alignment Scores:		2,21e-10	Length:	678
Pred. No.:	Score:	169.50	Matches:	51
Percent Similarity:		54.00%	Conservative:	30
Best Local Similarity:		74.00%	Mismatches:	54
Query Match:		11.84%	Indels:	15
DB:		14	Gaps:	5

US-09-905-247-1 (1-289) x CA840373 (1-678)

Qy	25	ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu	44
Db	211	GTGACATTCGACCTC-----GCCCAATCCGACACATATTCCAAGTTTCATTGCCGATCTC	264
Qy	45	ArgGlyAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSerLeu	64
Db	265	CAAACTGAACATAATTCCTCCGGCACAAACCGCATGTCAAATCCCGTGACACGGGCACCGCA	324
Qy	65	ProGlySer-----GlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThr	82
Db	325	ACCACATCCCAATGACCAACGATTCGTCTCTGCTCGACTCGAACAACACCTCGCAAAAACC	384
Qy	83	IleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArg-----	98
Db	385	ATCACCTCGCAATCGAGGTGACAGAGCTCTACGTCGTGCGCTATCGGCACAGTTCATAA	444
Qy	99	-----AlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGlu	113
Db	445	GAAGAAGGACGGACAAATGAAGGACGCGGCCAACCTTCCTCAAGATGCATCT---ACGGTT	501
Qy	114	AlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsn	133
Db	502	CGAAGGAGAAATCTTTTCATGGAGCAACAGTTCCG---AACCTTGCCTTCAAGGGACT	558
Qy	134	TyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuPro	153
Db	559	TACACATCTTTAGAGATGCTGCCAACCCACGACGAGAGCCATAGATTAGGGGTGGAT	618
Qy	154	AlaLeuAspSerAlaIleThrThreuPhe	163
Db	619	AAACTAGAGATTGGCATCGAGTCGGTTTAT	648

BE034055 811 bp mRNA linear EST 07-JUN-2000

MG04C05 MG Mesembryanthemum crystallinum cDNA 5' similar to

Ribosome-inactivating protein gelonin, mRNA sequence.

BE034055

BE034055 GI:8329064

EST.

SOURCE Mesembryanthemum crystallinum (common iceplant)

ORGANISM Mesembryanthemum crystallinum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Aizoaceae; Mesembryanthemum.

1 (bases 1 to 811)

Bohnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M.,

Pertea,H., Kawasaki,S., McColough,A., Michalowski,C.B.,

Palacios,C., Scarpa,G., Wheeler,M. and Zepeda,G.R.

REFERENCE

AUTHORS

Query Match:	11.84%	Indels:	15
DB:	14	Gaps:	5
US-09-905-247-1 (1-289) x CA838446 (1-667)			
QY	25	ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu	44
Db	209	GTGACATTCGACTC-----GCCAATCCGACACATATTCGAAGTTCATTGCGGATCTC	262
QY	45	ArgGlyAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuAlaArgSerSerLeu	64
Db	263	CAAACTGAACATAATTCCTCCGGCACAACCGCATGTCAAAATCCCGGTGACACGGGCAACCGCA	322
QY	65	ProGlySer-----GlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThr	82
Db	323	ACCACATCCCAATGACCAACGATTCGTCTGTGCACTCGAAACACCTCGCAAAAACC	382
QY	83	IleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArg-----	98
Db	383	ATCACCTCGCAATCGAGTGCACAGAGCTCTACGTCGTGCGCTATCGCACAGTTCAAA	442
QY	99	-----AlaGlyAspThrSerTyrPhePheAsnGlyAlaSerAlaThrGlu	113
Db	443	GAAGAAGCAGCGACAAATGAAGGACCGTGCCCAACTTCCTTCAAGATGCATCT---ACGGTT	499
QY	114	AlaAlaLysTyrValPheLysAspAlaMetArgLysValThrLeuProTyrSerGlyAsn	133
Db	500	GCGAAGGAGGAATCTCTTCATCGGAGCAACAGTTCGG---AACCTTCGCTTCAAGGGACT	556
QY	134	TyrGluArgLeuGlnThrAlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuPro	153
Db	557	TACACATCTTAGAAGATGCTGCRAACCAACGACGAGAACCATAGACTAGGGGTGGAT	616
QY	154	AlaLeuAspSerSerAlaIleThrThrLeuPhe	163
Db	617	AAACTAGAGTTTGGCATCGAGTCGGTTTAT	646
RESULT 14			
LOCUS	CA840373	678 bp	mRNA linear EST 12-DEC-2000
DEFINITION	MCT036G03 173841 Ice plant Lambda Uni-Zap XR expression library, 5 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase IV (5:30 PM). Mesembryanthemum crystallinum cDNA clone MCT036G03 5, mRNA sequence.		
ACCESSION	CA840373		
VERSION	CA840373.1	GI:26568138	
KEYWORDS	EST.		
SOURCE	Mesembryanthemum crystallinum (common iceplant)		
ORGANISM	Mesembryanthemum crystallinum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Aizoaceae; Mesembryanthemum.		
AUTHORS	1 (bases 1 to 678)		
TITLE	Cushman, J.C.		
JOURNAL	An expressed sequence tag database for the common ice plant,		
COMMENT	Mesembryanthemum crystallinum Unpublished (1997) Contact: Cushman JC Department of Biochemistry University of Nevada MS200, Reno, NV 89557-0014, USA Tel: 775-784-1918 Fax: 775-784-1650 Email: jcushman@unr.edu PCR Primers FORWARD: T3 20mer BACKWARD: T7 21mer Plate: 036 row: G column: 03 Seq primer: T3 20mer High quality sequence stop: 678. Location/Qualifiers 1. .678 /organism="Mesembryanthemum crystallinum"		
FEATURES	source		

TITLE Functional Genomics of Plant Stress Tolerance
JOURNAL Unpublished (2000)
COMMENT Contact: Michalowski, C.B.
University of Arizona
Bio Sciences West room 513, Tucson, AZ 85721, USA
Tel: 520-621-7982
Fax: 520-621-1697
Email: cbm@u.arizona.edu

Search completed: September 14, 2004, 21:56:37
Job time : 2484 secs

FEATURES An open reading frame exists.

Insert Length: 1 Std Error: 0.00.

source Location/Qualifiers
1..811
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
/tissue_type="roots"
/dev_stage="5-6 weeks"
/clone_lib="MG"
/note="grown in hydroponics, stress 400 mM NaCl (in 0.5
Hoagland's), 30 h stress"

ORIGIN

Alignment Scores:
Pred. No.: 2 97e-10 Length: 811
Score: 169.50 Matches: 59
Percent Similarity: 50.81% Conservative: 35
Best Local Similarity: 31.89% Mismatches: 69
Query Match: 11.84% Indels: 22
DB: 10 Gaps: 7

US-09-905-247-1 (1-289) x BE034055 (1-811)

```
QY 25 ValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPheIleSerAsnLeu 44
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
252 GTGACATTCAACCTC-----GCCCAACCCACACATATTCACATTTCATGACATCTCTC 305
QY 45 ArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeuArgSerSerLeu 64
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
306 CGAGCTCAACTCTGTGGCAACACCGCATGT---CAAATCCCGTGACACGGTCAACCGCA 362
QY 65 ProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGluThrIleSer 84
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
363 CCGGATCTCCAAGATTCTCTCTGTCGACCTCAAAACCAACCTCGCAAAAGACCATCACA 422
QY 85 ValAlaIleAspValThrAsnValTyrIleMetGlyTyrArg-----AlaGly 100
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
423 CTCGCAATCGAGTGCACAGCGTCTATGTCTGTGGCTACCGGACACAGCTTGGCGAANA 482
QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPheLys 120
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
483 GACCGTGCCAACTTCTTAGCGATGCT---CCACGGTCTGTAGGAACAACCTTTCAG 539
QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
540 GCGCGAGCGTTCGG---AACAATTCGTCGGAGGGATTACATAGCCCTAGAAAAAGCT 596
QY 141 AlaGlyLysIleArgGluAsnIle-ProLeuGlyLeuProAlaLeuAspSerAlaIleLeu 160
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
597 GCAAAGCAAGCGCGAATGCAATATGATGATTAGGTGATTAACCTAGATTGCGATCGA 656
QY 160 rThrLeuPheTyrTyrAsn-----AlaAsnSerAlaAlase 172
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
657 ATCCGTTTTTGGAGAGACCCCATTTATGTGTCAGGTTGAAGGCCAAATCTTGCTGATTG 716
QY 172 rAlaLeuMetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluG1 192
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
717 CTATACCAAGATTGGTCTC-----GAGCAACACGGTTTTTAGTTTTTGAAG 764
QY 192 nGlnIleGlyLys 196
Db |||:|||||
765 TACGGTACCCAAA 777
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 14, 2004, 21:15:17 ; Search time 455 Seconds
(without alignments)
3195.846 Million cell updates/sec

Title: US-09-905-247-1

Perfect score: 1432

Sequence: 1 MIRELVLSLLTLTLTPA.....AAMDDVPMVTFGCGSYAL 289

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0			
Ygapop 10.0			
Fgapop 6.0			
Delop 6.0			

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0905247/rnat_10092004_103936_4829/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPECL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=Blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0905247 @CGN 1 1 723 @rnat_10092004_103936_4829
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	/cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq*
2:	/cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq*
3:	/cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq*
4:	/cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq*
5:	/cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq*
6:	/cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq*
7:	/cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq*
8:	/cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq*
9:	/cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq*
10:	/cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq*
11:	/cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq*
12:	/cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq*
13:	/cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq2*
14:	/cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq*
15:	/cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq*
16:	/cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq*
17:	/cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq*
18:	/cgn2_6/ptodata/1/pubna/US10E_NEW_PUB.seq*
19:	/cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq*
20:	/cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					

ALIGNMENTS

RESULT 1
US-10-280-725B-3
; Sequence 3, Appli US/10-280-725B
; Publication No. US20040049025A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: Recombinant Viral Nucleic Acids
; FILE REFERENCE: LSBC-0109-US02
; CURRENT APPLICATION NUMBER: US/10/280,725B
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/557,941
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 08/484,341
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/923,692
; PRIOR FILING DATE: 1992-07-31
; PRIOR APPLICATION NUMBER: 07/600,244
; PRIOR FILING DATE: 1990-10-22
; PRIOR APPLICATION NUMBER: 07/641,617
; PRIOR FILING DATE: 1991-01-16
; PRIOR APPLICATION NUMBER: 07/737,899
; PRIOR FILING DATE: 1991-07-26
; PRIOR APPLICATION NUMBER: 07/739,143

1	1413	98.7	886	13	US-10-280-725B-3	Sequence 3, Appli
2	1413	98.7	886	15	US-10-280-679B-3	Sequence 3, Appli
3	728	50.8	913	15	US-10-127-890-13	Sequence 13, Appl
4	427	29.8	1855	16	US-10-394-511-31	Sequence 31, Appl
5	423	29.5	1698	9	US-09-785-921A-1	Sequence 1, Appli
6	423	29.5	1698	13	US-10-618-560-1	Sequence 1, Appli
7	423	29.5	1805	9	US-09-785-921A-10	Sequence 10, Appl
8	423	29.5	1805	13	US-10-618-560-10	Sequence 10, Appl
9	423	29.5	1855	16	US-10-394-511-3	Sequence 3, Appli
10	423	29.5	1855	16	US-10-394-511-5	Sequence 5, Appli
11	423	29.5	1855	16	US-10-394-511-7	Sequence 7, Appli
12	423	29.5	1855	16	US-10-394-511-9	Sequence 9, Appli
13	423	29.5	1855	16	US-10-394-511-13	Sequence 13, Appl
14	423	29.5	1855	16	US-10-394-511-15	Sequence 15, Appl
15	423	29.5	1855	16	US-10-394-511-17	Sequence 17, Appl
16	423	29.5	1855	16	US-10-394-511-19	Sequence 19, Appl
17	423	29.5	1855	16	US-10-394-511-21	Sequence 21, Appl
18	423	29.5	1855	16	US-10-394-511-23	Sequence 23, Appl
19	423	29.5	1855	16	US-10-394-511-25	Sequence 25, Appl
20	423	29.5	1855	16	US-10-394-511-27	Sequence 27, Appl
21	423	29.5	1855	16	US-10-394-511-29	Sequence 29, Appl
22	423	29.5	1855	16	US-10-394-511-33	Sequence 33, Appl
23	423	29.5	1855	16	US-10-394-511-35	Sequence 35, Appl
24	423	29.5	1855	16	US-10-394-511-37	Sequence 37, Appl
25	423	29.5	1855	16	US-10-394-511-39	Sequence 39, Appl
26	423	29.5	1855	16	US-10-394-511-48	Sequence 48, Appl
27	423	29.5	1855	16	US-10-394-511-50	Sequence 50, Appl
28	423	29.5	1855	16	US-10-394-511-52	Sequence 52, Appl
29	423	29.5	1855	16	US-10-394-511-54	Sequence 54, Appl
30	423	29.5	1855	16	US-10-394-511-74	Sequence 74, Appl
31	423	29.5	1855	16	US-10-394-511-77	Sequence 77, Appl
32	423	29.5	1855	16	US-10-394-511-80	Sequence 80, Appl
33	423	29.5	1855	16	US-10-394-511-83	Sequence 83, Appl
34	423	29.5	1855	16	US-10-394-511-86	Sequence 86, Appl
35	423	29.5	1855	16	US-10-394-511-89	Sequence 89, Appl
36	423	29.5	1855	16	US-10-394-511-92	Sequence 92, Appl
37	423	29.5	1855	16	US-10-394-511-95	Sequence 95, Appl
38	423	29.5	1855	16	US-10-394-511-98	Sequence 98, Appl
39	423	29.5	1855	16	US-10-394-511-101	Sequence 101, App
40	423	29.5	1855	16	US-10-394-511-104	Sequence 104, App
41	423	29.5	1855	16	US-10-394-511-110	Sequence 110, App
42	423	29.5	1855	16	US-10-394-511-113	Sequence 113, App
43	423	29.5	1855	16	US-10-394-511-116	Sequence 116, App
44	423	29.5	1855	16	US-10-394-511-119	Sequence 119, App
45	423	29.5	1855	16	US-10-394-511-122	Sequence 122, App

QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40
 DB 68 GTGAGGGCGATGTAGCTTCCTGTTATCAGGTGCAACAGCAGTTCCTATGAGGTTTC 127
 QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60
 DB 128 ATTTCAAACTGAGAAAAGCTCTTCCAAATGAAAGGAACTGTACGATATCCCTCTGTGA 187
 QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrSerTyrAlaAsp 80
 DB 188 CGTTCCTCTCTCCAGTCTTCACCGCTACGCAATGATCCATCTCAAAATTCAGCGCAT 247
 QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100
 DB 248 GAAACCATTTTCAGTGCCATAGACGTAACGACCTCTATATTTATGGATATCGCTGCG 307
 QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaLysTyrValPheLys 120
 DB 308 GATACATCTCTATTTTTCACGAGGCTTCTGCACAGAGCTGCAAAATATGTTATCAAA 367
 QY 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
 DB 368 GAGGCTATCGGAAAGCTTACGCTTCATATTCGGCAATTCGAAAGGCTTCAAACTGCT 427
 QY 141 AlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIleThr 160
 DB 428 GCGGCAAAATAGGGAATAATTCGCTTGGACTCCAGCTTGGACAGTGGCATTACC 487
 QY 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuIleGlnSer 180
 DB 488 ACTTGTCTTACTACACGCCAATCTCTGCTGGCTGGCACTTATGTTACTTCAATTCAGTCG 547
 QY 181 ThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIleGlyLysArgValAspLys 200
 DB 548 AGCTCTGAGGCTGCGAGGTATTAATTTATGACCAACAATTTGGGAAGCGCTTGACAAA 607
 QY 201 ThrPheLeuProSerLeuAlaIleSerLeuGluAsnSerTrpSerAlaLeuSerLys 220
 DB 608 ACCTTCTACCAAGTTAGCAATTAATAGTTTGGAAATAGTTGCTGCTCTCTCCAG 567
 QY 221 GlnIleGlnIleAlaSerThrAsnAsnGlyGlnPheGluSerProValValLeuIleAsn 240
 DB 668 CAATTCAGATAGCAGTACTAATATGACAGAGTTTGAACCTCTGTTGTGCTTATAAT 727
 QY 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnIle 260
 DB 728 GCTCAAAACCAACGATCATGATAACCAATGTTGATGCTGGAGTTGTAACTCCCAACATC 787
 QY 261 AlaLeuLeuLeuAsnArgAsnAsnMetAlaAlaMetAspAspValProMetThrGln 280
 DB 788 GCGTTGCTGCTGAATCGAAACAATATGCGCAGCATGGATGCGATGTTTCCATGACACAG 847
 QY 281 SerPheGlyCysGlySerTyrAlaLeu 289
 DB 848 AGCTTGGATGTGGAAGTTATGCTATT 874

RESULT 3

US-10-127-890-13

Sequence 13, Application US/10127890

Publication No. US20030166196A1

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Carroll, Stephen F.

Studnicka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESS: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-MAY-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-MAY-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 913 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-10-127-890-13

Alignment Scores:

Pred. No.: 5,558-86

Score: 728.00

Percent Similarity: 72.83%

Best Local Similarity: 58.87%

Query Match: 50.84%

DB: 15

Gaps: 2

US-09-905-247-1 (1-289) x US-10-127-890-13 (1-913)

QY 1 MetIleArgPheLeuValLeuSerLeuLeuLeuLeuThrLeuThrThrProAla 20

DB 10 ATGGTGAATTCCTTACTTCTTTTATTAATTCGCACTTCATTTGTTGTTCTACT 69

QY 21 ValGluGlyAspValSerPheArgLeuSerGlyAlaThrSerSerSerTyrGlyValPhe 40

DB 70 GCCAAGGCGATGTTAACTTCGATTTGTCGACTGCCACTGCAAAACCTCACAAAATTT 129

QY 41 IleSerAsnLeuArgLysAlaLeuProAsnGluArgLysLeuTyrAspIleProLeuLeu 60

DB 130 ATCGAAGATTCAGGCGCATCTTCCATTTAGCCATAAGTGTATGATATACCTCTACTG 189

QY 61 ArgSerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAsp 80

DB 190 TATTCACATTTCCGACTCCAGAGCTTCTATCTCCGATCTTCAAGTATGCGATAT 249

QY 81 GluThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGly 100

DB 250 GAAACCATCTCGGTGGCATAGATGTACCAACGTTTATGTTGTGGCGTATCGCACCGC 309

QY 101 AspThrSerTyrPhePheAsnGlyAlaSerAlaThrGluAlaLysTyrValPheLys 120

Db 310 GATGATCCACCTTTTAA---GAATCTCTCTCCCTGAAGCTATTAAACATCTTATCAAA 366
Qy 121 AspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThrAla 140
Db 367 --GGTACGCGGAAATACACTGCCATATACCGGTAAATATGAAATCTTCAACTGCT 423
Qy 141 AlaGlyLysLeuArgGluAsnLeuProLeuGlyLeuProAlaLeuAspSerAlaLeuThr 160
Db 424 GCACACAAAATAAGAGAGAAATATTGATCTTGGACTCCCTGCTTGGATGAGTAGGCCATTACC 483
Qy 161 ThrLeuPheTyrTyrAsnAlaAsnSerAlaAlaSerAlaLeuMetValLeuLeuGlnSer 180
Db 484 ACATGTTTTATACATGCCAATCTGCTCTCTGCAATGCTTGTACTAATCCAGAG 543
Qy 181 ThrSerGluAlaAlaArgTyrLysPheLeuGlnGlnLeuGlyLysArgValAspLys 200
Db 544 ACTGCAGAGCTGCAAGATTTAAGTATATCAGGACACGTTGCTAAGTATGTTGCCACT 603
Qy 201 ThrPheLeuProSerLeuAlaLeuLeuSerLeuGluAsnSerTyrSerAlaLeuSerLys 220
Db 604 AACTTTAAGCCAAATCTAGCCATCATAGCTTGGAAATCAATGCTGCTCTCTCCAA- 662
Qy 221 GlnLeuGlnLeuAlaSerThrAsnAsmGlyGlnPheGluSerProValValLeuLeuAsn 240
Db 663 CAATCTTTTGGCGCAGAAATCAAGGAGGAAATTTAGAAATCTCTGCACTTATAAA 722
Qy 241 AlaGlnAsnGlnArgValThrIleThrAsnValAspAlaGlyValValThrSerAsnLeu 260
Db 723 CCTACCGGGAAACGGTTTCAAGTAACCAATGTTGATTCAGATGTTGTAAGGTAATATC 782
Qy 261 AlaLeuLeuLeuAsn 265
Db 783 AACTCTGCTGTAAC 797

RESULT 4

US-10-394-511-31
; Sequence 31, Application US/10394511
; Publication No. US20040009551A1
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; TITLE OF INVENTION: Viral or Parasitic Infections
; FILE REFERENCE: 10447-27
; CURRENT APPLICATION NUMBER: US/10/394,511
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/063,715
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: US 60/045,148
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 1855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pAP-242 insert
US-10-394-511-31

Alignment Scores:
Pred. No.: 1.64e-45 Length: 1855
Score: 427.00 Matches: 97
Conservative: 58.76%
Best Local Similarity: 35.40% Mismatches: 93
Query Match: 29.82% Indels: 20
DB: 16 Gaps: 8

US-09-905-247-1 (1-289) x US-10-394-511-31 (1-1855)

Qy 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheLeuSerAsnLeu 44

Db 136 ATAACTTTTACCACGCGGTGCCACTGTGCAAAAGCTACAAACATTTTATCAGAGCTGTT 195
Qy 45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspLeuProLeuLeu-----Arg 61
Db 196 CGCGTGTGTTTAACTGAGCTGATGTGAGACATGATATACAGTGTGTCGCAAAACAGA 255
Qy 62 SerSerLeuProGlySerGlnArgTyrAlaLeuLeuHleHisLeuThrAsnTyrAlaAspGlu 81
Db 256 GTTGTGTTTGCCTATAAACCAACGCTTTATTATTAGTTGAACCTCTCAAAATCATGAGAGCTT 315
Qy 82 ThrIleSerValAlaLeuAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101
Db 316 TCTGTATCATAGCGCTGGATGTACCAATGCATATGTGTGCGCTACCGTGTGCGAAAT 375
Qy 102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119
Db 376 AGGCATATTTCTTTCATCTGCAATCAGAAAGATCGAAGCAATCCTCATCTCTTTC 435
Qy 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
Db 436 ACTGATGTTCAAAATCGATATACATTCGCCCTTTGGTGGTAAATATGATAGTACTTGAACAA 495
Qy 140 AlaAlaGlyLysLeuArgGluAsnLeuProLeuProAlaLeuAspSerAlaLeu 159
Db 496 CTTGCTGTGTAATCTGAGAGAAATATCGAGTTGGAAATGTGCTCCACTAGAGAGGCTATC 555
Qy 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174
Db 556 TCAGCGCTTTTATTATTACAGTACTGGTGGCACTCAGCTTCCAACTCTGCTCGTCTCTTT 615
Qy 175 MetValLeuLeuGlnSerThrSerGluAlaAlaArgTyrLysPheLeuGlnGlnLeu 194
Db 616 ATAAATTTGCATCCAAATGATTTTCAGAGCAGCAAGATTTCCAAATATATTGAGGGGAAATG 675
Qy 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaLeuLeuSerLeuGlu 212
Db 676 CGCAGAGAAATAGGTACACCGGAGATCTGCACAGATCTTAGCGTAAATACACTTGAG 735
Qy 213 AsnSerTyrSerAlaLeuSerLysGlnLeuGlnLeuAlaSerThrAsnAsnGlyGlnPhe 232
Db 736 AATAGTTGGGGAGACTTTCCACTGCAATTCAA-----GAGTCTAACCAAGAGGCTTT 789
Qy 233 GluSerProValValLeuLeuAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db 790 GCTAGTCCAAATTCAACTGCAAGAGCGTAAATGGTTTCCAAATTCAGTGTG-----TACGAT 843
Qy 253 AlaGlyValValThrSerAsnLeuAlaLeuLeuLeuAsnArg-----AsnAsn 268
Db 844 GTGAGTATATTAATCCCTATCATAGTCTCATGGTGTATAGATGCGACCTCCACCATCG 903
Qy 269 MetAlaAlaMetAspAspValProMetThrGlnSerPhe 282
Db 904 TCACAGTTGTTTTCGAGAACTATCCAAATAGTGCAAAATTTT 945

RESULT 5

US-09-785-921A-1
; Sequence 1, Application US/09785921A
; Patent No. US2002009434A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: LIT-PI-529
; CURRENT APPLICATION NUMBER: US/09/785,921A
; CURRENT FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 1
; LENGTH: 1698
; TYPE: DNA
; ORGANISM: Ricinus communis
US-09-785-921A-1

```
Alignment Scores:
Pred. No.: 4,8e-45 Length: 1698
Score: 423.00 Matches: 94
Percent Similarity: 60.24% Conservatives: 59
Best Local Similarity: 37.01% Mismatches: 85
Query Match: 29.54% Indels: 16
DB: 9 Gaps: 7

US-09-905-247-1 (1-289) x US-09-785-921A-1 (1-1698)
QY 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyArgValPheIleSerAsnLeu 44
Db 97 ATAAACTTTTACCACAGCGGGTCCACTGTGCAAGCTACACAACTTTATCAGAGCTGT 156
QY 45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg 61
Db 157 CGCGGTGCGTTTAACTGCTGATGTCACCAATGTCATATGTCGCTACCATGTTGCCAAACAGA 216
QY 62 SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu 81
Db 217 GTTGTGTTTGCCTATAACCAACCGTTTATTTAGTTGAACCTCTCAATCATCGAGCTT 276
QY 82 ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101
Db 277 TCTGTATACATAGCCCTGATGTCACCAATGTCATATGTCGCTACCGTCTGGAAT 336
QY 102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119
Db 337 AGCGCATATTTCTTTCATCCTGACATCAGGAAGATGCAAGCAATCACTCATCTTTTC 396
QY 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgGlnThr 139
Db 397 ACTGATGTTCAAAATCGATATACATTCGCTTGTGTTGTAATATATGATAGACTTGAACA 456
QY 140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
Db 457 CTTGCTGTAATCTGAGAGAAATATCGAGTTGGGAAATGGTCCACTAGAGAGGCTATC 516
QY 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174
Db 517 TCAGCGCTTTATTTATACAGTACTGGTGGCACTCAGCTTCCAACTCTGCTCTCTTT 576
QY 175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnIle 194
Db 577 ATAATTTGCATCCAAATGATTTTCAGAGCAGCAAGATTCATATATTTGAGGAGAAATG 636
QY 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 212
Db 637 CGCAGGAAATTAGGTACAAACCGAGATCTGCACAGATCTCTAGCGTAATTTACACTTGAG 696
QY 213 AsnSerTrpSerAlaLeuSerLysGlnIleAlaSerThrAsnGlyGlnPhe 232
Db 697 AATAGTTGGGGAGACTTTTCCACTGCAATTCAA-----GAGTCTAACCAAGGAGCCTTT 750
QY 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db 751 GCTAGTCCAAATTCACCTGCAAGAGCGTAATGTTTCCAAATTCAGTGTG-----TACGAT 804
QY 253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuAsnArg 266
Db 805 GTGAGTATATTAATCCCTATCATGCTCTCATGTGTATAGA 846

RESULT 6
US-10-618-560-1
; Sequence 1, Application US/10618560
; Publication No. US20040048784A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; FILE REFERENCE: LIT-PI-529
```



```

Db      558 TCAGCGCTTTATTATTACAGTACTGGTGGCACTCAGCTTCCAACTCGGTGCTGCTTCCTTT 617
QY      175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIle 194
Db      618 ATAATTGGCATCCAAATGATTTTCAGAGACGACGAGATTCCTCAATATATTGGGAGAAATG 677
QY      195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 212
Db      678 CGCAGAGAAATTAGGTACAAACCGAGATCTGCACCATCGATCTCGGTAATTACACTTGAG 737
QY      213 AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe 232
Db      738 AATAGTTGGGGAGACTTCACTCACTGCAATTCAA-----GAGTCTAACCAAGAGCGCTTT 791
QY      233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db      792 GCTAGTCCCAATTCACTGCAAGACGTAATGTTCCAAATTCAGTGTG-----TAGCAT 845
QY      253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuAsnArg-----AsnAsn 268
Db      846 GTGAGTATATTATTCCTCATCATAGCTCTCATGCTGTATAGATCGGCACCTCCACCATCG 905
QY      269 MetAlaAlaMetAspAspValProMetThrGlnSerPhe 282
Db      906 TCACAGTTTGTCTCTCAAACTACCACTCACTCAAAATTTT 947

RESULT 9
US-10-394-511-3
; Sequence 3, Application US/10394511
; Publication No. US20040009551A1
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; FILE OF INVENTION: Viral or Parasitic Infections
; FILE REFERENCE: 10447-27
; CURRENT APPLICATION NUMBER: US/10/394,511
; CURRENT FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/063,715
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: US 60/045,148
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cathepsin B linker regions of pAP-214
US-10-394-511-3

Alignment Scores:
Pred. No.: 5,58e-45 Length: 1855
Score: 423.00 Matches: 94
Percent Similarity: 60.24% Conserv: 59
Best Local Similarity: 37.01% Mismatches: 85
Query Match: 29.54% Indels: 16
DB: 16 Gaps: 7

US-09-905-247-1 (1-289) x US-10-394-511-3 (1-1855)
QY      25 ValSerPheArgLeuSerGlyAlaThrSerSerThrGlyValPheIleSerAsnLeu 44
Db      136 ATAACTTTACACAGCGGTGGCACTGTGAAGACTACAACTTTATCAGAGCTGTT 195
QY      45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg 61
Db      196 CGCGTCTGTTTAAACAACTGGAGCTGATGAGACATGATATACCATGTTGCCAAACAGA 255
QY      62 SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu 81

```

```

Db      256 GTTGGTTTGCTATAAACCAACGGTTTATTATTAGTCAACTCTCAATCATCGAGAGCTT 315
QY      82 ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101
Db      316 TCTGTTACATATAGCGTGGATGTACCAATGTCATATGTGGTCGCTACCGTGTGGAAAT 375
QY      102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119
Db      376 AGGCAATTTCTTTCATCTGCAATTCAGGAAGATGCAGAGCAATCACTCATCTTTTC 435
QY      120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
Db      436 ACTGATGTTCAAAATCGATATACATTCGCTTGGTGGTAAATTATGATACACTTGAACAA 495
QY      140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
Db      496 CTTGCTGGTAATCTGAGAGAAATATCGAGTTGGGAAATGGTCCACTAGAGGAGGCTATC 555
QY      160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174
Db      556 TCAGCGCTTTATTATTACAGTACTGGTGGCACTCAGCTTCCAACTCTGGCTGCTCTT 615
QY      175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIle 194
Db      616 ATAATTTCATCCAAATGATTTTCAGAGCAGCAGAGATTCCTCAATATATTGAGGAGAAATG 675
QY      195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 212
Db      676 CGCAGAGAAATTAGGTACAAACCGAGATCTGCACAGATCTTAGCGTAATTACACTTGAG 735
QY      213 AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe 232
Db      736 AATAGTTGGGGAGACTTTCACACTGCAATTCAA-----GAGTCTAACCAAGAGCTTT 789
QY      233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db      790 GCTAGTCCCAATTCACTGCAAGACGTAATGTTCCAAATTCAGTGTG-----TAGCAT 843
QY      253 AlaGlyValValThrSerAsnIleAlaLeuLeuLeuLeuLeuAsnArg 266
Db      844 GTGAGTATATTAACTCCCTATCATAGCTCTCATGCTGTATAGA 885

RESULT 10
US-10-394-511-5
; Sequence 5, Application US/10394511
; Publication No. US20040009551A1
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; FILE OF INVENTION: Viral or Parasitic Infections
; FILE REFERENCE: 10447-27
; CURRENT APPLICATION NUMBER: US/10/394,511
; CURRENT FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/063,715
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: US 60/045,148
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pAP-216 insert
US-10-394-511-5

Alignment Scores:
Pred. No.: 5,58e-45 Length: 1855
Score: 423.00 Matches: 94
Percent Similarity: 60.24% Conserv: 59

```

```

Best Local Similarity: 37.01% Mismatches: 85
Query Match: 29.54% Indels: 16
DB: 16 Gaps: 7

US-09-905-247-1 (1-289) x US-10-394-511-5 (1-1855)

Qy 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu 44
Db 136 ATAACTTTTACCAAGCGGGTCCACTGTGCAAGCTACACAACTTTATCAGAGCTGT 195

Qy 45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg 61
Db 196 CGCGTCTTTTAACTGAGCTGATGTGAGATGATATACCACTGTTCGCAACACAGA 255

Qy 62 SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu 81
Db 256 GTTGCTTGCCTATAAACCAGCGTTTATTAGTTGAACCTCAAAATCATCGAGAGCTT 315

Qy 82 ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101
Db 316 TCTGTTACTAGCCCTGAGTGCACCAATGCATATGTGGTGGTCCACTAGAGGAGCTATC 435

Qy 102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119
Db 376 AGCGCATATTTCTTCATCCTGACAATCAGGAAGATGCAAGCAATCACTCATCTTTTC 435

Qy 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
Db 436 ACTGATGTTCAAATCGATATACATTCGCTTTGGTGGTAAATATATGATAGACTTGAACA 495

Qy 140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
Db 496 CTTGCTGGGAGAGCTTTCCACTGCAATTCAA-----GAGTCTAACCAAGAGGCTTT 789

Qy 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaIleSerLeuLeu 174
Db 556 TCAGCGCTTTATTATTACAGTACTGTGTGGCACTTCAGCTTCCAACTCGTCTCTTT 615

Qy 175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnIle 194
Db 616 ATAATTTGCAATTCAGAGCAAAATATCGAGTTGGGAATGGTCCACTAGAGGAGCTATC 675

Qy 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 212
Db 676 CGCAGAGAAATAGGTACACCGAGATCTGCACAGATCTTAGCGTAATTACACTTGAG 735

Qy 213 AsnSerTrpSerAlaLeuSerLysGlnIleAlaSerThrAsnGlnPhe 232
Db 736 AATAGTTGGGAGAGCTTTCCACTGCAATTCAA-----GAGTCTAACCAAGAGGCTTT 789

Qy 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db 790 GCTAGTCCAATTCAGTCAAGAGCAAGTAAATGGTTCCAAATTCAGTGTG-----TAGCAT 843

Qy 253 AlaGlyValValThrSerAsnIleAlaLeuLeuAsnArg 266
Db 844 GTGAGTATATATCCCTATCATAGCTCTCATGCTGTATAGA 885

; PRIOR APPLICATION NUMBER: US 60/045,148
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pAP-218 insert
US-10-394-511-7

Alignment Scores: 5.58e-45 Length: 1855
Pred. No.: 423.00 Matches: 94
Score: 60.24% Conservative: 59
Percent Similarity: 37.01% Mismatches: 85
Best Local Similarity: 29.54% Indels: 16
Query Match: 16 Gaps: 7

US-09-905-247-1 (1-289) x US-10-394-511-7 (1-1855)

Qy 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu 44
Db 136 ATAACTTTTACCAAGCGGGTCCACTGTGCAAGCTACACAACTTTATCAGAGCTGT 195

Qy 45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg 61
Db 196 CGCGTCTTTTAACTGAGCTGATGTGAGATGATATACCACTGTTCGCAACACAGA 255

Qy 62 SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu 81
Db 256 GTTGCTTGCCTATAAACCAGCGTTTATTAGTTGAACCTCAAAATCATCGAGAGCTT 315

Qy 82 ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101
Db 316 TCTGTTACTAGCCCTGAGTGCACCAATGCATATGTGGTGGTCCACTAGAGGAGCTATC 435

Qy 102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119
Db 376 AGCGCATATTTCTTCATCCTGACAATCAGGAAGATGCAAGCAATCACTCATCTTTTC 435

Qy 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
Db 436 ACTGATGTTCAAATCGATATACATTCGCTTTGGTGGTAAATATATGATAGACTTGAACA 495

Qy 140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
Db 496 CTTGCTGGGAGAGCTTTCCACTGCAATTCAA-----GAGTCTAACCAAGAGGCTATC 555

Qy 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaIleSerLeuLeu 174
Db 556 TCAGCGCTTTATTATTACAGTACTGTGTGGCACTTCAGCTTCCAACTCGTCTCTTT 615

Qy 175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnIle 194
Db 616 ATAATTTGCAATTCAGAGCAAAATATCGAGTTGGGAATGGTCCACTAGAGGAGCTATC 675

Qy 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu 212
Db 676 CGCAGAGAAATAGGTACACCGAGATCTGCACAGATCTTAGCGTAATTACACTTGAG 735

Qy 213 AsnSerTrpSerAlaLeuSerLysGlnIleAlaSerThrAsnGlnPhe 232
Db 736 AATAGTTGGGAGAGCTTTCCACTGCAATTCAA-----GAGTCTAACCAAGAGGCTTT 789

Qy 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db 790 GCTAGTCCAATTCAGTCAAGAGCAAGTAAATGGTTCCAAATTCAGTGTG-----TAGCAT 843

Qy 253 AlaGlyValValThrSerAsnIleAlaLeuLeuAsnArg 266
Db 844 GTGAGTATATATCCCTATCATAGCTCTCATGCTGTATAGA 885

; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; TITLE OF INVENTION: Viral or Parasitic Infections
; FILE REFERENCE: 10447-27
; CURRENT APPLICATION NUMBER: US/10/394,511
; CURRENT FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/063,715
; PRIOR FILING DATE: 1997-10-29

```

[illegible]

QY 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
Db 436 ACTGATGTTCAAAATCGATATATATCGCTTGGTGAATATAGTAGACTTTGAACAA 495
QY 140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
Db 496 CTTGCTGGTAACTCGAGAGAAATATCGAGTTGGGAATGTCCACATAGAGGAGCTATC 555
QY 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174
Db 556 TCAGCGCTTATTATATACAGTACTGGTGGCACTCAGCTTCCAACTCTCGCTCGTTCCTTT 615
QY 175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnIle 194
Db 616 ATAATTTGCATCCAAATGATTTCAGAACGACGAAGATTCCAAATATATTGGGGAGAAATG 675
QY 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleSerLeuGlu 212
Db 676 CGCAGAGAAATAGTACACCGGAGATCTGCACAGATCTTAGCGTAAATACACTTGAG 735
QY 213 AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe 232
Db 736 AATAGTTGGGGAGACTTTCCACTGCAATTCAA-----GAGTCTTAACCAAGGAGCCCTTT 789
QY 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db 790 GCTAGTCCAAATCAACTGCAAGAGCGTAATGGTTCCAAATTCAGTGTG-----TACGAT 843

RESULT 14

US-10-394-511-15
; Sequence 15, Application US/10394511
; Publication No. US20040009551A1
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; FILE REFERENCE: 10447-27
; CURRENT APPLICATION NUMBER: US/10/394,511
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/063,715
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: US 60/045,148
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pap-226 insert
US-10-394-511-15

Alignment Scores:
Pred. No.: 5,58e-45 Length: 1855
Score: 423.00 Matches: 94
Percent Similarity: 60.24% Conservative: 59
Best Local Similarity: 37.01% Mismatches: 85
Query Match: 29.54% Indels: 16
DB: 16 Gaps: 7

US-09-905-247-1 (1-289) x US-10-394-511-15 (1-1855)

QY 25 ValSerPheArgLeuSerGlyAlaThrSerSerTyrGlyValPheIleSerAsnLeu 44
Db 136 ATAACACTTACCAGCGGTGCCACTGTGCAAAAGCTACACAACTTTATCAGAGCTGTT 195

QY 45 ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeuLeu-----Arg 61
Db 196 CGCGGTGTTTAAACAACCTGAGCTGATGTGAGACATGATATACCACTGTTGCCAAACAGA 255
QY 62 SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu 81
Db 256 GTTGGTTGGCTATAAAACCAACGCTTTATTATTAGTTGAACCTCTCAAAATCATCGAGAGCTT 315
QY 82 ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp 101
Db 316 TCTGTTACATTAGCGCTGGATGTCCCAATCATATGTGTCGCTACCGTGGCTGGAAT 375
QY 102 ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe 119
Db 376 AGCGCATATTCTTTCATCTCGTACCAATCAGGAAGATGCAGAAACAATCACTCATCTTTTC 435
QY 120 LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr 139
Db 436 ACTGATGTTCAAAATCGATATATACATTCGCCCTTGGTGAATATATAGTAGACTTTGAACAA 495
QY 140 AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle 159
Db 496 CTTGCTGGTAACTCGAGAGAAATATCGAGTTGGGAATGTCCACATAGAGGAGCTATC 555
QY 160 ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu 174
Db 556 TCAGCGCTTATTATATACAGTACTGGTGGCACTCAGCTTCCAACTCTCGCTCGTTCCTTT 615
QY 175 MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnIle 194
Db 616 ATAATTTGCATCCAAATGATTTCAGAACGACGAAGATTCCAAATATATTGGGGAGAAATG 675
QY 195 GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleSerLeuGlu 212
Db 676 CGCAGAGAAATAGTACACCGGAGATCTGCACAGATCTTAGCGTAAATACACTTGAG 735
QY 213 AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyGlnPhe 232
Db 736 AATAGTTGGGGAGACTTTCCACTGCAATTCAA-----GAGTCTTAACCAAGGAGCCCTTT 789
QY 233 GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp 252
Db 790 GCTAGTCCAAATCAACTGCAAGAGCGTAATGGTTCCAAATTCAGTGTG-----TACGAT 843

RESULT 15

US-10-394-511-17
; Sequence 17, Application US/10394511
; Publication No. US20040009551A1
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; FILE REFERENCE: 10447-27
; CURRENT APPLICATION NUMBER: US/10/394,511
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/063,715
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: US 60/045,148
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1855
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: PAP-228 insert
US-10-394-511-17

Alignment Scores:

Pred. No.:	5.58e-45	Length:	1855
Score:	423.00	Matches:	94
Percent Similarity:	60.24%	Conservative:	59
Best Local Similarity:	37.01%	Mismatches:	85
Query Match:	29.54%	Indels:	16
DB:	16	Gaps:	7

US-09-905-247-1 (1-289) x US-10-394-511-17 (1-1855)

QY	25	ValSerPheArgLeuSerGlyAlaThrSerSerSeryrGlyValPheIleSerAsnLeu	44
Db	136	ATAAACTTTACCAACAGCGGGTGCACCTGTGCAAGACTACACAAACTTTTATCAGAGCTGTT	195
QY	45	ArgLysAlaLeuProAsnGluArgLysLeu---TyrAspIleProLeu-----Arg	61
Db	196	CGCGTTCGTTTAAACAACCTGGAGCTGATGTGAGACATGATATACCAAGTGTGTCCAACAGCA	255
QY	62	SerSerLeuProGlySerGlnArgTyrAlaLeuIleHisLeuThrAsnTyrAlaAspGlu	81
Db	256	GTTCGGTTTGCCTATAAAACCAACGGTTTATTATTAGTTGAACCTCTCAAAATCATGCAGAGCTT	315
QY	82	ThrIleSerValAlaIleAspValThrAsnValTyrIleMetGlyTyrArgAlaGlyAsp	101
Db	316	TCGTGTTACATTTAGCGCTGGATGTCCACCAATCATATGTGGTCGGCTACCGTCTGGAAT	375
QY	102	ThrSerTyrPhePhe-----AsnGlyAlaSerAlaThrGluAlaAlaLysTyrValPhe	119
Db	376	AGGCATATTCTTTCATCTCTGCAATCAGGAAGATGCAGAAAGCAATCACTCATCTTTTC	435
QY	120	LysAspAlaMetArgLysValThrLeuProTyrSerGlyAsnTyrGluArgLeuGlnThr	139
Db	436	ACTGATGTCAAAATTCGATATACATTCGCCCTTTGGTGGTAATTATGATAGACTTGAACAA	495
QY	140	AlaAlaGlyLysIleArgGluAsnIleProLeuGlyLeuProAlaLeuAspSerAlaIle	159
Db	496	CTTGCTCGTGAATCTGAGAGAAAATATCGAGTTGGGAAATGTCCTACATAGAGAGGCTATC	555
QY	160	ThrThrLeuPheTyrTyrAsnAlaAsn-----SerAlaAlaSerAlaLeu	174
Db	556	TCACGCGTTTATTATTACAGTACTGGTGGCACTCAGCTTCCAATCTGGCTCGTTCCTTT	615
QY	175	MetValLeuIleGlnSerThrSerGluAlaAlaArgTyrLysPheIleGluGlnGlnIle	194
Db	616	ATAATTGGCATCCAAGTGAITTCAGAGCAGCAAGATCCCAATATATTCAGGGAGAAATG	675
QY	195	GlyLysArgVal-----AspLysThrPheLeuProSerLeuAlaIleIleSerLeuGlu	212
Db	676	CGCAGCAGAGATAGGTACAAACCGGAGATCTGCACCAAGATCCTAGCGTAAATTACACTTGG	735
QY	213	AsnSerTrpSerAlaLeuSerLysGlnIleGlnIleAlaSerThrAsnAsnGlyClnPhe	232
Db	736	AATAGTTGGGGAGACTTTCACCTGCCAATTCAA-----GAGTCTAACAGGAGCGCTTT	789
QY	233	GluSerProValValLeuIleAsnAlaGlnAsnGlnArgValThrIleThrAsnValAsp	252
Db	790	GCTPAGTCCAATTCAACTGCAGACAGCGTAATGGTTCCAAATTCAGTGTG-----TAGCAT	843
QY	253	AlaGlyValValThrSerAsnIleAlaLeuLeuLeuAsnArg	266
Db	844	GTGAGTATATTAATCCCTATCATAGCTCTCATGGTGTATAGA	885

Search completed: September 14, 2004, 23:07:40
Job time : 463 secs

